Self-organization from structural refrigeration

Eric Smith

Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, USA

The self-organization of a classical current is studied, in an exactly solvable model where both the quantum statistics over microhistories of particles, and the macroscopic phenomenology, can be computed in closed form. It is shown that for thermodynamically reversible systems, the Jaynes formulation of statistical mechanics naturally extends to include explicit macroscopic dynamics and heterogeneities in temperature, while preserving the structure of partition functions, effective potentials, and ground states of the equilibrium theory. Self-organization in such reversible systems is constrained by entropy transport through engine and refrigeration cycles, rather than by diffusion in gradients. Limitations in the ability to decompose such systems sensibly into components with additive entropies, and in the extrapolation of entropy functions from equilibrium forms, are discussed with examples.

I. INTRODUCTION

A. Untangling self-organization from dissipation

"Self-organization"—the dynamical emergence of temporally or spatially ordered macrostates from microphysics subjected to less-ordered boundary conditions—has become an actively investigated phenomenon, especially in chemistry [1,2], granular flow [3,4], and the fine structure of friction [5–7]. In all these studies, when ordering depends on the discreteness of the particles or events, the analysis is carried out in an irreversible quasistatic limit, so that evolution is a Markov process with respect to states characterized by their configuration (as opposed to current) variables [3,5]. When ordering is a result of mass action, it is typically analyzed in terms of phenomenological state variables and equations of motion, whose forms are locally those of equilibrium systems (Ref. [2], Chap. 3). Temporal ordering emerges from these classical equations, regarded as dynamical systems evaluated far from their equilibrium solutions. Since all of these analyses apply to extreme-dissipative limits, there is a tendency to conflate self-organization (SO) with dissipative structures [8], self-organized criticality [9], or some alternative [10] but similarly irreversible process.

This paper argues that the criteria for SO are informational in nature, and that they can arise in the dynamics of ideally reversible systems. The reversibility considered here is not a tautological consequence of a finetuned microscopic reversibility (as in Ref. [11]), but rather the standard, macroscopic reversibility of Carnot's theorem in thermodynamics [12]. Here not dissipation, but the partitioning of entropy flows among elementary engine and refrigeration cycles, emerges as the organizing constraint on the system. Thus the association of dynamical organization with dissipation is at best case dependent.

The idea that SO can sensibly arise from reversible dynamics is demonstrated here with mass-action models with ordinary macroscopic state variables. The key difference in this analysis from the works cited above [1,2,8], though, is the way in which order is defined from the residual uncertainties in microhistories given macrohistories, by direct extension of the treatment of time-independent states in equilibrium statistical mechanics [13]. The measure of uncertainty is the usual Shannon/Boltzmann entropy of a density matrix, but the evolution of this density is studied in the reversible, rather than the Markovian, limit. The point emphasized is the key role that projection of the description of a system into its components has, in creating component entropies that can have any nontrivial dynamics. These *necessarily coarse-grained* entropies [14] induce measures of both the degree and the complexity of the dynamically generated order. The concepts of both entropy production and entropy transport are derived as consequences of the interaction between Hamiltonian evolution and the coarse graining implied by component projection—they are not taken as primitives.

The entropies that result do not, in general, have the phenomenological forms of equilibrium, though by construction they have the same function as measures of residual uncertainty, extended to apply to dynamical macrostates. Though at leading order the entropies of components are additive as in the well-known chemical mass-action models [2], it is the subleading orders in the evolving whole-system entropy that quantify the spontaneously generated organization. As expected from the careful analysis of Maxwell demons [15,16], only when these entropies of macroscopic organization are included as corrections to the transported entropies of classical Carnot's theorem, is the correct formulation of the second law of thermodynamics, and in particular its reversible limit, respected.

The difference between the true entropies of history, and their corresponding equilibrium phenomenological forms, serves to emphasize how the quantitative understanding of SO depends on the proper definition of classical states and macroscopic descriptions of transformations in statistical systems. It is well known that in comparisons of equilibrium ensembles, only coarse-grained entropies *can* increase, and they are only assured to increase relative to some reference description, if the state variables in the reference represent the correct and *complete* constraints on the otherwise-maximum-entropy distribution [14]. Macroscopic averages that account only for a subset of the complete set of state variables of some ensemble define an intrinsically *coarser*

ensemble, and the second law says nothing about how subsequent characterizations of the true ensemble may relate to such a loose initial characterization. The mere existence of macroscopic averages does not imply that they are adequate as state variables, or place any constraint on the evolution of entropy functions computed from them.

The same is true for ensembles whose macroscopic averages evolve with some interesting time dependence, except that the set of eligible state variables becomes larger. Thus even the formulation of the second law is predicated on the nontrivial technical problem of identifying which measurements (available at or after any elapsed time) represent the true constraints on the uncertainty of histories. It is shown here that, in general, the state variables of a self-organizing system will include currents or explicitly time-dependent histories, as well as the static configurations adequate to characterize equilibrium. A restriction to instantaneous densities, or other such equilibrium forms, results in a loose characterization of the true state, and loss of the ability to apply the second law as any tight constraint on their evolution. The errors in estimating the entropy, arising simply from the restriction to an inappropriate and incomplete set of state variables, can be quite sufficient to completely hide the actual entropy changes associated with the formation of order.

B. Mode of analysis and results

The current paper is not an axiomatization of SO even for reversible systems, nor does it aim to define a completely general notation to capture all aspects of what one may wish to call organization. Good general works relating to measures of order and complexity in Markov and general statistical systems are Refs. [17,18].

Instead, this paper presents a rather thorough analysis of an exactly solvable example, which is at once an elementary quantum statistical ensemble, and a familiar instance of dynamical formation of order at the classical level. The example incorporates matter flow through degrees of freedom naturally interpreted as system and environment components. It incorporates heterogeneous and perpetually time-dependent temperatures, and yet is completely described within Jaynes's formulation of "equilibrium" statistical mechanics [19]. It therefore demonstrates the subtleties of identifying classical state variables when measurements may be bounded in time, and the proper entropy function of those, in a case where both are known to exist and can be expressed in closed form. It demonstrates the (provable) need for nonstatic classical variables to define the ensemble of microhistories, and shows quantitatively what errors are made by loose descriptions restricted to the phenomenology of equilibrium. The entropies associated with both formation of order, and simple heat flow, occur with distinct powers in a regular small-parameter expansion, and both the classical Carnot theorem and its corrections from subsystem organization can be derived exactly. Finally, the model admits an effective potential description of the classical evolution, and makes contact with a similar treatment of heterogeneous and dynamical temperatures carried out previously with finite-temperature field theory methods [20,21].

The model chosen as example is a linear quantum-oscillator idealization of the spontaneous formation of an inductor (L) current from a capacitor (C) charge separation in a resistanceless LC circuit, or any number of other equivalent processes of macroscopic oscillation between potential and kinetic order, ultimately created by interference of quantum coherent states. While this is the simplest nontrivial model possible, it captures the process by which nontrivial self-organization happens in empirically interesting cases, among which the self-organizing reversible thermoacoustic engines [22] have already been analyzed as dynamical critical systems [21]. Furthermore, the Carnot organization of macroscopic entropy transport (once that has been derived as a process) has clear and relevant isomorphisms in biochemical processes such as enzyme-catalyzed polymerization. Therefore the form of the model has been chosen to maximize the explicit overlap with these cases.

The analysis leads to two main conclusions of a fundamental nature.

The Jaynsian "maximum-ignorance" formulation of statistical mechanics [19] offers a principled definition of classical states, which is general enough to include the classically reversible systems on an equal footing with those in static equilibrium. Properly understood, it follows more or less from definition that classical reversibility, and not merely equilibrium, is the proper domain of "equilibrium" statistical mechanics.

The familiar thermodynamic potentials, such as the Helmholtz or Gibbs potentials [13], are simply the leading terms in an expansion in heterogeneities of the environmental constraints on an ensemble. Where the equilibrium form captures the static response of the system to the homogeneous constraints imposed by the environment, the succeeding terms represent the coupling to heterogeneities, and the time-dependent or current responses of the system to them. This will be demonstrated explicitly in the example, but it is important enough to deserve a separate abstract treatment in the following section.

Though limited in scope, this discussion serves as a first step in a systematic analysis of SO in its own terms. The reversible limit considered here readily admits the steady dynamical incorporation of uncertainty by tracing over unobserved aspects of the environmental state or interactions. A precise analysis of the origins of irreversibility, and its consequences for the forms of possible ordered states, can thus be pursued directly. These extensions will be performed in future work.

C. Layout of the paper

The paper is organized as follows. Section II abstracts the generalization of thermal effective potentials to reversible systems with heterogeneous temperatures and possibly explicit dynamics. Section III reviews the relevant features of the empirically well-studied thermoacoustic engines, and the considerations from biochemistry relevant to reversible SO. Section IV introduces the actual model, and computes its properties as a microcanonical ensemble, a system of self-organizing components, and a classical engine. Section V de-

fines the proper entropy of histories, and the generalization of the Helmholtz potential for the dynamical model. Finally, Sec. VI discusses consequences for the interpretation of complexity and the notion of ground states.

Following the main text, six appendixes establish notation and introduce the tools and supporting calculations that relate this derivation to other analyses of SO in engines [20,21]. They are written to be readable as an independent track on mathematical methods, and to introduce a number of tools for handling heterogeneous-temperature partition functions that have not been used before. Of particular methodological interest are Appendix C, which introduces Gaussian-coherent representation of thermal ensembles, and Appendix F, which derives the relation of the current discrete method for handling heterogeneous temperature, to finite-temperature field-theory methods developed previously [20].

II. EFFECTIVE POTENTIALS FOR HETEROGENEOUS-TEMPERATURE SYSTEMS

A. Constraints, distributions, and state variables

An obvious and often-advanced goal [3] (Ref. [2], Sec. 7.8) in the analysis of dynamical SO is to relate the formation of order to processes like spontaneous symmetry breaking [23] in equilibrium critical phenomena, which have a well-understood conceptual foundation. The most compact, and still one of the conceptually clearest, tools for understanding the relation of expressed to hidden symmetries is the effective potential [24], which has indeed motivated an industry of deriving potential methods for stochastic (Markov) processes [25,26].

An advantage of studying SO first in reversible systems is that, even when their order is explicitly dynamical, these yield effective-potential descriptions that are a direct continuation of the familiar thermodynamic potentials of equilibrium. One can therefore introduce the effects of environments that impose heterogeneous temperatures (or other thermodynamic potentials) and induce flows, without simultaneously taking on the additional difficulties of breaking time-reversal symmetry and having to reinterpret the meanings of operators in the Markov case [25] relative to those in the underlying Hamiltonian field theory.

The usual Helmholtz potential A for a static system in equilibrium thermal contact with an external world at homogeneous temperature $T \equiv 1/\beta$ is defined by

$$\beta A = \beta U - S. \qquad (1)$$

The structure of A directly portrays its function. One may regard U (average internal energy) and S (Shannon/Boltzmann entropy) as properties of whatever distribution represents the system, well defined whether or not it is coupled to an environment. When coupling is introduced, U is the system property on which the environment imposes a constraint, and S is the system property maximized when average internal energy is the *only* constraint. β is the constant of proportionality, characterizing the environment, which determines the slope dS/dU to which the system settles. Only in the maximum-entropy configuration does the

scalar constraint on dS/dU determine an entire distribution, at which point U and T become *state variables* sufficient to define the resulting ensemble.

In equilibrium theory, S is assumed to be a function of static configuration variables only. If there is order in the state that maximizes S at constrained U, even if it arises by symmetry breaking and is not "selected" by the fine structure in the environment, one still thinks of the need for order as being "imposed" by the environmental constraint. That is, the system is as disordered as it can be when coupled to that environment.

Self-organizing systems are anomalous with respect to the representation given by the equilibrium effective potential, because they are typically more ordered than any of its entropy maxima. The excess order only arises, though, when heterogeneities in the environment induce flows through the system. In the simplest case, the characterization equivalent to Eq. (1) must therefore contain at least some set of temperatures $\{T_i \equiv 1/\beta_i\}$. The potential A cannot, in general, identify the response of the system to such an environment, and the minimal set of state variables needed to correctly compute the residual microstate uncertainty is no longer static. There can, however, be a function $\mathcal{A}$ that generalizes A appropriately. If, from the $\{\beta_i\}$ (by whatever prescription) one computes some average inverse temperature $\bar{\beta}$ that couples to U, $\mathcal{A}$ will, in general, take the form

$$\bar{\beta}\mathcal{A} = \bar{\beta}U + \sum_{(i,j)} (\beta_i - \beta_j)J_{ij} - S. \qquad (2)$$

The $\{J_{ij}\}$ are some set of time-dependent state variables effectively constrained by the corresponding temperature differences $\beta_i - \beta_j$. They may represent classical histories of reversible systems, in which case the $\beta_i - \beta_j$ will generally be data on some Cauchy surface, which specify the initial conditions for the ensemble in question. Alternatively, in irreversible systems, the $\{J_{ij}\}$ may represent currents with time-invariant values, but which break time-reversal symmetry, and the $\beta_i - \beta_j$ may represent steady-state boundary conditions. (Derivation of this latter form will be provided in future work incorporating irreversibility.) S in Eq. (2) will, in general, depend on U and all of the J_{ij}. As in the equilibrium case, the β_i will select U and J_{ij} values through the derivatives of S. At this maximum of the more general entropy, the distribution again depends only on the U and J_{ij} values, and it becomes sensible to interpret them as the state variables.

The model introduced in Sec. IV generates an explicit instance of Eq. (2), in which the dynamical system has an exact representation as a density matrix ρ over quantum states of simple-harmonic oscillators, and S has a closed-form evaluation as the trace of $\rho \ln \rho$. Particle transport through classical simple-harmonic oscillation is the only nonequilibrium form of order, and $\beta_i - \beta_j$ are initial temperature boundary conditions in two "reservoirs" of the environment, which then imprint themselves forever on oscillating charge/current history variables J_{ij}.

B. Classical reversibility is generalized "equilibrium"

The definition of *state variable* used here is: one of a set of constraints intrinsic to what one calls a thermodynamic system. It then follows that entropies computed from the state variables must be intrinsic properties of the system, and not artifacts of a deliberate coarseness in its characterization. This is in slight contrast to the point of view usually adopted in statistical mechanics [2,13], that the state variables are the only macroscopic quantities *worth* measuring, even to characterize a unique instance of a strongly chaotic system. In the latter use, the very definition of entropy is context dependent [14], upon what one considers worth measuring.

One can acknowledge the context dependence of entropies computed from macroscopic averages, while at the same time allowing them to be intrinsic properties of the system, by associating classical state variables with the *preparation conditions* that admit a class of instances, which one later measures as an ensemble. (An ideal example is the Stern-Gerlach experiment.)

From such a definition, there follows an unambiguous distinction between reversible and irreversible processes. Since every ensemble admits an arbitrarily large collection of measurements, and since measurements in the real world require either lapse of time or selection of parallel instances, the statistical distinction to be made is whether an arbitrary set of measurements is predicated on a fixed uncertainty of histories, given once and for all, or not. Thus, one may define the following.

(1) A process described by an ensemble is *reversible* if an arbitrary collection of measurements on the ensemble is constrained by a fixed uncertainty over microscopically specified histories.

(2) A process is *irreversible* if the specification of evolution in the ensemble admits a changing set of microhistories, and thus the order and time requirements of measurements affect the sets of histories on which they are made.

Under this definition, it is a tautology that the expansion of which Eq. (2) lists the first few possible terms is general for classically reversible processes. Hamiltonian dynamics implies that the indexing of microhistories at any time has an isomorphism to an indexing at any other time. Macroscopic reversibility implies that the uncertainty of the ensemble is fixed; hence there is an isomorphism between any pair of complete sets of state variables, such as those bounded at or after two different times. The technical challenge is simply to identify the isomorphism. Thus, the proper domain of Jaynes's statistical mechanics is the category of classically reversible processes, and not just the static equilibria within it.

The question of when the state variables of either a reversible or irreversible system, bounded by some sequence of times, retain a criterion of *simplicity* (such as time locality or spatial smoothness) is a separate and technical issue. It demands that the state variables both capture the constraints from preparation, and satisfy an ongoing restriction of form. When both requirements are satisfied, one may be sloppy about distinguishing what is intrinsic to the ensemble from what are the limitations of the experimenter, and gloss over the distinction emphasized here. The special place of static
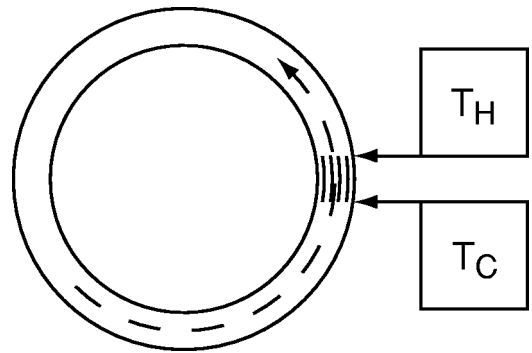

FIG. 1. The idealized traveling-wave thermoacoustic engine. Squares are reservoirs at temperatures T_H and T_C, and the concentric circles represent the resonator. Short parallel lines in the resonator are the stack of closely spaced plates, coupled to the reservoirs at either end as indicated by the arrows. Dashed arrow indicates the direction of propagation of the spontaneously generated traveling wave.

equilibria among the reversible processes is that they more readily permit this identification.

III. MOTIVATIONS FOR THE MODELS

Two aspects of the model constructed below are constrained to make contact with empirical cases of SO. First, the simple redistribution of quantum bosons to form a classical current is used to represent the formation of order in a real, reversible, dynamical critical system. For this the representative instance is the traveling-wave thermoacoustic engine. Second, it is asserted that organization according to Carnot cycles is a plausible principle for the formation of order. The map to enzymatic catalysis in biochemistry, while it requires a conversion from thermal to chemical heterogeneity, is an interesting as well as conceptually important domain in which to argue that this is the case.

A. The self-starting, self-organizing heat engines

There is a class of acoustic, ideal-gas heat engines [22] with the following structure and properties. An engine consists of a periodic (typically toroidal) resonator filled with an ideal gas, and a stack of plates or pins in the flowstream of the gas, with one end of each plate coupled to a source of heat, and the other to a sink, as shown in Fig. 1. Thermalized phonons may be regarded as the only excitations in the system.

When the temperature difference between the ends of the plates exceeds a critical value, a classical traveling wave at the fundamental frequency, and arbitrarily determined phase, spontaneously condenses from the thermal phonon bath. In the ideal inviscid limit of the gas, the critical temperatures are $0^{\pm}$ for the (orthogonal) opposite-sense waves, traveling in the direction of the temperature gradient. The traveling wave implements an ideal (Carnot-efficient) Stirling cycle [27], from which the work extracted is stored in increased in-phase magnitude of the existing wave, by an acoustic process akin to Dicke superradiance [28]. Significant fractions

of this ideal efficiency have been achieved in realized engines [29].

If small dissipation is introduced in the classical gain equation as a regulator, via a Langevin equation, the critical temperatures for the dual modes separate to nonzero positive and negative values, and the spontaneously formed sound saturates as $\sqrt{\Delta T - \Delta T_C}$, where ΔT and ΔT_C are, respectively, the imposed and critical temperature differences on plate ends. This behavior can be derived from an effective field theory for the fundamental modes [21] of the Onsager-Machlup form [30].

The engines are self-organizing, by the criterion that the classical-wave configuration has lower entropy than an isoenergetic thermal population of the fundamental wave vector states. Furthermore, this SO process has been explicitly mapped to an equilibrium critical phenomenon [20,21] with an extension of finite-temperature field theory methods to incorporate heterogeneous temperatures. [The correspondence of that method to Eq. (2) is developed in Appendix F.] Yet the organized state itself is nothing more than an asymmetric population of wave vector states for thermalized bosons. Formation of just such a current state can be reproduced in the appropriate linear oscillator model.

B. Chemical engine cycles in biology

A shared feature of the most commonly treated mass-action self-organizing systems, such as Benard convection [31], the Belousov-Zhabotinsky reaction [32], and the irreversible (standing-wave) thermoacoustic engines [33], is that organization depends intrinsically on diffusion in gradients. Taking these systems as models for the origin of biochemical order [2] presumes the associated intrinsic inefficiency of diffusion, in one form or another, as requisite to ordering.

Even apart from the empirical observation that many biochemical processes achieve remarkable fractions of ideal efficiency [34], the surprisingly mechanical sequence through which fundamental ordering operations such as enzymatically catalyzed polymerization take place suggests that a description in terms of chemical engines powering chemical refrigerators may be more appropriate. While this paper is not the place to justify such a description in detail, to appreciate the applications of the model it is useful to understand in what sense polymerization is abstractly just another refrigeration process.

The decomposition of the simplest polymerization model has two pairs of reservoirs. One pair is the cytoplasmic solution versus the formed polymer as distinct reservoirs for monomers. The different entropies of arrangement for a monomer in either reservoir gives an entropic contribution to the chemical potential difference. The other pair is pyrophosphate versus orthophosphate as alternative reservoir states for phosphate groups, with ATP or GTP generally acting as the stabilizing reservoir for the energetic pyrophosphate bond. The enzyme is a mechanical mediator that allows the phosphates to decay only in lockstep with incorporation of the monomers into the polymerized form [41].

The essential point to appreciate is that in reversible SO, the entropy flux behaves like any classical Noether current
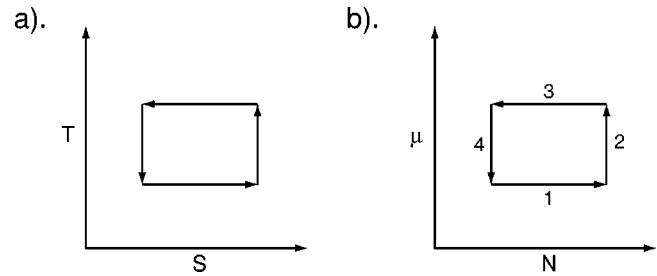

FIG. 2. The Carnot refrigeration cycle (a), and its chemical equivalent (b). T is the temperature of the working volume (and of any reservoir to which it momentarily couples), and S is its entropy. Alternatively, N is the number of some particle species conducted through the chemical refrigerator, and μ is the chemical potential of any reservoir with which it would instantaneously be in equilibrium. Numbers on the legs of the chemical cycle correspond to the stages in enzymatic polymerization detailed in the text.

[20], and thus is not mathematically distinguished from a flux of particle number (monomer or phosphate group) across chemical potentials. Figure 2 shows the correspondence between the thermal (temperature/entropy) Carnot cycle, and its chemical (chemical potential/particle number) counterpart. The legs in the chemical cycle correspond directly to the simplified steps in polymerization, as follows.

(1) A monomer diffuses from the cytoplasm to the active site of an enzyme, where an n-mer is already bound. This step is the intake of the monomer from a reservoir with which the enzyme is in equilibrium.

(2) Whether at binding to the enzyme, or through activation with a phosphate group, or through conformational change of the enzyme powered at the release of the last cycle, the monomer is given free energy in excess of the bound form in the polymer. This step effectively changes the chemical potential of the monomer to its new, higher value, in isolation from the (unactivated) cytoplasmic population.

(3) The enzyme, in lowering the energy barrier to binding, allows the monomer to hop into the bound form. Possibly in some combination with the next step, this is the rejection of the monomer to the higher-energy reservoir.

(4) Finally the now $(n+1)$-mer is advanced along the active region of the enzyme, whose configuration is then returned to one that can accept another monomer from solution. Conformational (elastic) energy may or may not be added to the enzyme in this step, which in either case causes the return of the monomer-binding site to the potential of the solution.

IV. MODELING ENGINE AND RESERVOIR DYNAMICS: THE BLOCH CRYSTAL ENGINE

A linear model illustrating the phenomena discussed up to this point is shown in Fig. 3. It is built from an even number M of linear, quantum harmonic oscillators with identical level spacing. Two oscillators are considered to be the engine, and the remaining $M-2$ are collected into symmetric "left" and "right" reservoirs. The reservoirs have a literal interpretation as plates of a capacitor, and the resonator as an inductor, in a dilute-Fermi-gas limit where fermionic statis-
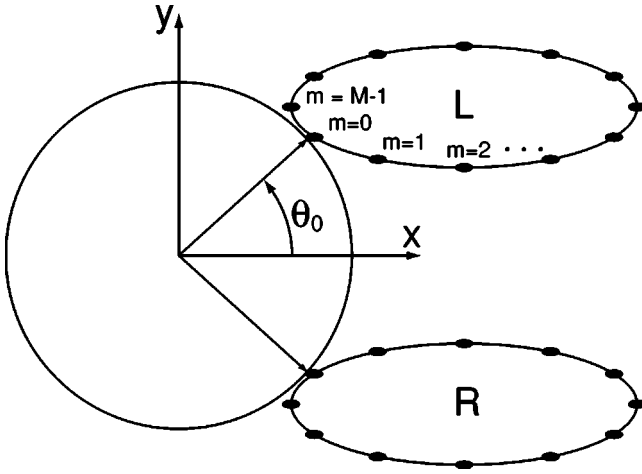
FIG. 3. Diagram of the Bloch-crystal engine, to be compared with Fig. 1. Big circle corresponds to the resonator; its two excitations are the fundamental-mode phonons of the low-energy effective theory. Axes of symmetry for spatial standing waves are marked x and y, and axes of spatial standing waves coupled to the reservoirs are at angles θ_0 and $-\theta_0$ relative to x. Spatial modes in each reservoir are indexed $m = 0, \ldots, M^{L,R} - 1$ in each of left and right sectors. There is a number-exchange coupling between standing waves in the resonator and the $m = 0$ spatial oscillator of the corresponding reservoir.

tics of electrons can be ignored. Definitions and notation for the quantum harmonic oscillator are reviewed in Appendix A, along with a number of manipulations of scalar and vector operators that will be used in the following constructions. Definitions of coarse grainings, and the forms that will be applied to this system, are given in Appendix B.

Elementary excitations along orthogonal axes x and y in the engine are created by two raising operators $a_x^\dagger$ and $a_y^\dagger$, respectively. These correspond to the (second-quantized) low-energy effective description of the thermoacoustic engine in Ref. [21]. The Hermitian conjugate operators to $a_x^\dagger$ and $a_y^\dagger$ are called a_x and a_y, and the excitation number operators are $\hat{n}_x$ and $\hat{n}_y$. In terms of these, the free Hamiltonian for the engine is just its total excitation number

$$H_0^E \equiv \hat{n}_x + \hat{n}_y. \qquad (3)$$

The reservoirs are given slightly more structure, as Bloch crystals, with the nearest-neighbor Hamiltonian

$$H_0^L \equiv \sum_{m=0}^{M^L - 1} \hat{n}_m^L - \frac{\gamma}{2} [a_m^{\dagger L} a_{m-1}^L + a_m^{\dagger L} a_{m+1}^L] \qquad (4)$$

for the left reservoir, and similarly for $L \to R$. m is periodically identified, and $M^L = M^R$, so that $M = 2M^L + 2$.

The creation operator for a normalized wave vector state in the L reservoir is

$$a_k^{\dagger L} \equiv \frac{1}{\sqrt{M^L}} \sum_{m=0}^{M^L - 1} e^{-ikm} a_m^{\dagger L}, \qquad (5)$$

so that in terms of the associated number operators

$$H_0^L \equiv \sum_k \hat{n}_k^L (1 - \gamma \cos k), \qquad (6)$$

with sum over $k \in (0, \ldots, M^L - 1) \times 2\pi/M^L$, and similarly for $L \to R$.

The coupling of the engine to the reservoirs is a simplified version of the coupling used for real thermoacoustic engines, although the latter, like lasers, have a nonlinear gain equation. As noted in Sec. III, since the information represented in the current itself is of interest, the form of the gain equation leading to it may be changed in this case without losing the effect. A convenient form comes from the interaction Hamiltonian

$$H_{\text{int}} = -\frac{g}{\sqrt{2}\cos\theta_0} [a_{\theta_0}^\dagger a_0^L + a_{-\theta_0}^\dagger a_0^R + \text{H.c.}]$$

$$= -g[a_x^\dagger a_0^S + \tan\theta_0 a_y^\dagger a_0^A + \text{H.c.}], \qquad (7)$$

where the standing-wave excitations at angles $\pm\theta_0$ are created by the operators

$$a_{\theta_0}^\dagger \equiv \cos\theta_0 a_x^\dagger + \sin\theta_0 a_y^\dagger \qquad (8)$$

and

$$a_{-\theta_0}^\dagger \equiv \cos\theta_0 a_x^\dagger - \sin\theta_0 a_y^\dagger, \qquad (9)$$

respectively. Subscript 0 on the reservoir operators denotes *spatial* index $m = 0$, since $k = 0$ would make the Bloch ring pointless. g is the coupling strength of the engine to the reservoirs, and may be taken small or of order unity, as desired.

The operators in the first line of Eq. (7) do not have orthogonal canonical commutation relations at general θ_0, so the second line gives an expansion in operators that do, with symmetric and antisymmetric lowering operators in the reservoirs defined as

$$a_0^S \equiv \frac{a_0^L + a_0^R}{\sqrt{2}} \qquad (10)$$

and

$$a_0^A \equiv \frac{a_0^L - a_0^R}{\sqrt{2}}. \qquad (11)$$

Everywhere H.c. denotes Hermitian conjugate of the terms that appear explicitly.

The closed engine/reservoir system evolves microscopically under the Hamiltonian

$$H = H_0^E + H_0^L + H_0^R + H_{\text{int}}. \qquad (12)$$

There is a spatial basis $\{x, y, m^L, m^R\}$, which naturally decomposes into components on which heterogeneous environmental couplings can be imposed. (Alternatively one could

use $\{x, y, k^L, k^R\}$.) The excitations in these bases differ from excitations in the eigenstate basis by a unitary transformation of the raising and lowering operators, and this property leads to nontrivial flow of particles among the spatial or wave number projections.

A. Preparation of heterogeneous thermal initial conditions

A convenient separation of scales can be achieved, between the fundamental oscillator frequency (set to one in these units), and the frequency of particle exchange, by taking the angle θ_0 very close to $\pi/4$ in the operating state. It will be assumed below that $\theta_0 < \pi/4$, as in Fig. 3.

This coupling admits a very convenient way to impose heterogeneous thermal initial conditions, because at $\theta_0 = \pi/4$, the engine operators are orthogonal, with canonical commutation relations. Therefore one can imagine starting with a "preparation coupling"

$$H_{\text{int, prep}} = -\frac{g}{\sqrt{2}\cos\theta_0}[a_l^\dagger a_0^L + a_r^\dagger a_0^R + \text{H.c.}]$$

$$= -\frac{g}{\sqrt{2}\cos\theta_0}[a_x^\dagger a_0^S + a_y^\dagger a_0^A + \text{H.c.}], \qquad (13)$$

where $a_l^\dagger$ and $a_r^\dagger$ are defined as in Eq. (8) and Eq. (9), respectively, except with $\theta_0 \to \pi/4$. The entire left and right sectors will then decouple, each can be prepared in a thermal state at an independently specified temperature, and the density matrix for the ensemble will then be the product of densities for the two sectors. Coupling can then be introduced with perturbative strength $\nu \equiv g(1 - \tan\theta_0)$, by simply rotating the reservoir contact points slightly toward the x direction.

The easiest case in which to understand the origin of particle transport, and indeed the only case where time-local measurements will lead to reversible dynamics, is $M^L = M^R = 1$. Letting the Bloch exchange coupling $\gamma \to 0$, the whole preparation Hamiltonian can be written in matrix form as

$$H_{\text{prep}} = [a_l^\dagger \ a_0^{L\dagger}]\begin{bmatrix} 1 & -g/\sqrt{2}\cos\theta_0 \\ -g/\sqrt{2}\cos\theta_0 & 1 \end{bmatrix}\begin{bmatrix} a_l \\ a_0^L \end{bmatrix}$$

$$+ [a_r^\dagger \ a_0^{R\dagger}]\begin{bmatrix} 1 & -g/\sqrt{2}\cos\theta_0 \\ -g/\sqrt{2}\cos\theta_0 & 1 \end{bmatrix}\begin{bmatrix} a_r \\ a_0^R \end{bmatrix}. \qquad (14)$$

Left and right eigenstate excitations are manifestly created by operators $(a_l^\dagger \pm a_0^{L\dagger})/\sqrt{2}$ and $(a_r^\dagger \pm a_0^{R\dagger})/\sqrt{2}$, both with eigenvalues $1 \mp g/\sqrt{2}\cos\theta_0$.

It is shown in Appendix C that thermal densities can be written in terms of Gaussian integrals over coherent states, and the notation K is introduced for the kernel matrix of the Gaussian integral. The eigenvalues of K for homogeneous thermal densities are the inverses of the mean occupation numbers in the eigenstate basis. In the left and right preparation bases, these take the approximate forms at high temperature,

$$K_\pm^L \equiv e^{\beta^L(1 \mp g/\sqrt{2}\cos\theta_0)} - 1 \to \beta^L(1 \mp g/\sqrt{2}\cos\theta_0) \quad (15)$$

and

$$K_\pm^R \equiv e^{\beta^R(1 \mp g/\sqrt{2}\cos\theta_0)} - 1 \to \beta^R(1 \mp g/\sqrt{2}\cos\theta_0). \quad (16)$$

If $g/\sqrt{2}\cos\theta_0$ is chosen close to one, the population in each sector can be dominated by the symmetric (low-frequency) eigenstate to any desired degree, reducing the analysis of this problem to that of single-particle thermal states. After all the properties of that limit are understood, the case of general g, etc., can be examined.

Not only x and y excitation numbers, but also those in l and r and various traveling-wave bases will be of interest in the analysis that follows. Therefore it is useful to remark that all of these bases differ from each other only by unitary transformations of the creation operators, and it is shown in Appendix D that the marginal distributions of any Gaussian coherent-state densities are exactly thermal in any such bases. Further, the expected mean excitation numbers are just the diagonal elements of K^{-1} in the corresponding representation [Eq. (D1)], so it is useful to define mean-excitation matrices

$$\bar{n}^L \equiv (K^L)^{-1} \qquad (17)$$

and

$$\bar{n}^R \equiv (K^R)^{-1}, \qquad (18)$$

from which all the engine excitation numbers can be derived by unitary transformation.

B. Single mode-driven oscillations

It was actually possible to diagonalize the preparation Hamiltonian in either left and right sectors, or in the basis $\{x, S_0, y, A_0\}$. However, the thermal initial conditions of interest are only diagonal in the left/right basis, whereas the eigenstates with $\theta_0 < \pi/4$ require diagonalization in $\{x, S_0, y, A_0\}$. The matrix representation of the dynamical Hamiltonian is

$$H = [a_x^\dagger \ a_0^{S\dagger}]\begin{bmatrix} 1 & -g \\ -g & 1 \end{bmatrix}\begin{bmatrix} a_x \\ a_0^S \end{bmatrix} + [a_y^\dagger \ a_0^{A\dagger}]$$

$$\times \begin{bmatrix} 1 & -g\tan\theta_0 \\ -g\tan\theta_0 & 1 \end{bmatrix}\begin{bmatrix} a_y \\ a_0^A \end{bmatrix}, \qquad (19)$$

and its eigenstate excitations are created by $(a_x^\dagger \pm a_0^{S\dagger})/\sqrt{2}$, with eigenvalue $1 \mp g$, and $(a_y^\dagger \pm a_0^{Y\dagger})/\sqrt{2}$, with eigenvalue $1 \mp g\tan\theta_0$.

No difficulty is incurred because the basis in which the initial thermal projection factors is different from the eigenstate basis. Appendix E shows that, when the coherent-state representation of thermal density matrices is used for L and R sectors, the density which is their product simply defines an M-dimensional kernel K, in which the factor matrices K^L and K^R become diagonal blocks [Eq. (E6)]. The partition

function has a basis-independent definition, from which the expected excitation numbers at any time are easily extracted by suitable similarity transform of K or its inverse, $\bar{n}$.

The $t=0$ transformation from $\{l,L_0,r,R_0\}$ to $\{x,S_0,y,A_0\}$ operators puts the mean number matrix in the representation

$$\bar{n}_0 \equiv \frac{1}{2}\begin{bmatrix} \bar{n}^L+\bar{n}^R & \bar{n}^L-\bar{n}^R \\ \bar{n}^L-\bar{n}^R & \bar{n}^L+\bar{n}^R \end{bmatrix}, \qquad (20)$$

where the blocks $\bar{n}^L$ and $\bar{n}^R$ are defined by Eq. (17) and Eq. (18), respectively.

Since the eigenstates advance their phases according to the eigenvalues of Eq. (19), $\bar{n}$ evolves by similarity transform with the diagonal time-evolution operator. Because of the interaction Hamiltonian used to define the preparation of the sectors, the full eigenstates are superpositions of the sector eigenstates with the same engine/reservoir symmetry or asymmetry. In other words, in the eigenbasis, the block factors $\bar{n}^L$ and $\bar{n}^R$ are themselves diagonal, and the only non-identity contribution to the similarity transform comes from the energy difference matrix

$$E_A - E_S = \nu\begin{bmatrix} 1 & \\ & -1 \end{bmatrix}. \qquad (21)$$

Supposing that the coupling between the L and R sectors was turned on at $t=0$, the number matrix at time t, called $\bar{n}_t$, is just

$$\bar{n}_t \equiv \frac{1}{2}\begin{bmatrix} \bar{n}^L+\bar{n}^R & (\bar{n}^L-\bar{n}^R)e^{-i(E_A-E_S)t} \\ (\bar{n}^L-\bar{n}^R)e^{i(E_A-E_S)t} & \bar{n}^L+\bar{n}^R \end{bmatrix}. \qquad (22)$$

It is clear that under transformation back to $\{l,L_0,r,R_0\}$, the mean l and r number densities oscillate with frequency ν between the values induced by the two initially imposed temperatures. The x and y mean numbers meanwhile remain constant at the average of l and r means.

C. Bases for entropy accounting

The $M^{L,R}=1$ model is intermediate between the simplicity of an equilibrium system, and the complexity of the $M^{L,R}>1$ models, which though formally reversible by the definition given in Sec. II, have classical state variables that are difficult to extract from late-time measurements, and whose definitions do not readily generalize beyond this example. At $M^{L,R}=1$, one can see why restricting to equilibrium state variables is inadequate, while keeping the true set of state variables time local (though time dependent), so that they have a natural description as indices to macrohistories.

It is shown in Appendix D that the marginal distributions for which matrix (22) gives mean numbers are all exactly thermal, in any excitation basis, at any t. Each marginal has a well-defined effective temperature, with the expected relation to its mean occupation number. Thus temperatures of modes contain all of the information restricting the ensemble

at any time. They are also, of course, redundant, since a well-defined set of temperatures exists in any basis for $\bar{n}_t$ related by unitary transformation to Eq. (22).

The question of interest is when these temperatures admit a projection of this coupled system onto a product of engine and reservoir components, in which classical properties of the components are sufficient to specify the constraints on the whole distribution. If there is always such a projection, one is free to restrict to equilibrium state variables, and assign them equations of motion as in the dissipative models [2]. Here one can show explicitly that, in terms of the "charge densities" $\bar{n}_l,\bar{n}_r,\bar{n}_x,\bar{n}_y,\bar{n}_m^L,\bar{n}_m^R$, that will not be possible. It will be possible to project into "system" and "environment" components, but in the system projection one must admit as state variables *current* densities as well.

The entropy of the marginal distribution of any identically thermal mode is defined in terms of the mean occupation number in Eq. (C6). The condition that occupation numbers measured in a projection onto components be proper state variables is equivalent to the condition that the sums of entropies computed from them equal the true entropy of the distribution over microhistories. It is clear why sums of marginal entropies need not be conserved in general. The operation of factoring into components (replacing joint distributions with products of marginals) is a coarse graining, which can lose information about system/environment correlations.

The condition of thermal reversibility of a factored system, taken as a commonplace in equilibrium systems, is seen in dynamical systems to place a delicate requirement on the choice of state variables, even when the set of microhistories is unchanged over time. It is a condition that a sum of entropies after marginalization be conserved in time, as the dynamics changes the relation of the distribution to the component projection.

In the example here, the marginals do recover a conservation law as $g/\sqrt{2}\cos\theta_0$ is taken close to one. In that limit, the low-frequency states account for essentially all of the population in any basis, and there is a *redundancy* of the information contained in the reservoir $m=0$ and the engine states. It then becomes possible to use reservoir marginal distributions as *proxies* for part of the order in the engine, allowing engine bases to be explored to account for any additional order not measurable in the factored reservoir marginals alone. Note that this limit is not necessitated by anything fundamental; it is used to compensate for the prejudice that, because the marginals are thermal, the constraints can be inferred from their temperatures alone.

Quantum mechanically, standing-wave and traveling wave excitations are not independent in an engine with only two degrees of freedom. However, classically, there may be information represented in the mean excitation number in one basis, which is only contained in the joint probabilities across different excitation numbers in the other basis. Thus exploring the various number bases in the engine will be the key to recovering an approximate description of the complete constraints on the distribution of interfering microstates. The necessary excitations beyond those already considered are created by the operators

$$a_+^\dagger = \frac{a_x^\dagger - i a_y^\dagger}{\sqrt{2}} \quad (23)$$

and

$$a_-^\dagger = \frac{-i a_x^\dagger + a_y^\dagger}{\sqrt{2}}. \quad (24)$$

It follows from the commutation relations that all of

$$\bar{n}_x + \bar{n}_y = \bar{n}_l + \bar{n}_r = \bar{n}_0^L + \bar{n}_0^R = \bar{n}_+ + \bar{n}_- \quad (25)$$

are equal and time independent, and from Eq. (22) that $\bar{n}_x - \bar{n}_y \equiv 0$. Unitary transformation to the l, r, L_0, R_0, or from these to the $+,-$ bases at any time, gives

$$(\bar{n}_l - \bar{n}_r)_t \approx (\bar{n}_l - \bar{n}_r)_0 \cos \nu t \quad (26)$$

and

$$(\bar{n}_+ - \bar{n}_-)_t \approx -(\bar{n}_l - \bar{n}_r)_0 \sin \nu t \quad (27)$$

from Eq. (22). Furthermore, because the populations of both engine and reservoir L and R modes come entirely from the interference of the same pair of states, $(\bar{n}_0^L - \bar{n}_0^R)_t \approx (\bar{n}_l - \bar{n}_r)_t$. This is how the choice of coupling makes $\bar{n}_0^L$ and $\bar{n}_0^R$ informational proxies for the l and r coherence in the engine.

Note that $\bar{n}_+ - \bar{n}_-$, from an equilibrium point of view, is a *current* variable, not a charge variable. It is necessary in Eq. (27) to track the constraints on initial temperatures at times $\nu t = (2j+1)\pi$ for integer j, and yet is not within the set of equilibrium densities used as default for state variables.

At high temperature, the thermal entropy of Eq. (C6) (for any one component) is asymptotically equal to the logarithm of the mean excitation number, so to second order in fluctuation amplitudes,

$$\frac{d}{dt}[S(\bar{n}_0^L) + S(\bar{n}_0^R)] \neq 0, \quad (28)$$

and also

$$\frac{d}{dt}[S(\bar{n}_+) + S(\bar{n}_-)] \neq 0, \quad (29)$$

though total excitation numbers in the two sectors are preserved independently. The phase offset between the two oscillations, however, implies that to the same order

$$\frac{d}{dt}[S(\bar{n}_0^L) + S(\bar{n}_0^R) + S(\bar{n}_+) + S(\bar{n}_-)] \approx 0. \quad (30)$$

Under almost any weakening of the restrictions in this example, the basis used here will no longer yield a reversible description, and the informational entropies will be obscured by a larger background of fluctuations due to inadequacies of the coarse graining. However, with some sensitivity to the

structure of the dynamics, it will generally be possible to preserve factorability if one can use time-nonlocal measurements. First, though, having understood the origin of a need for current-state variables, it is of interest to extract the classical and self-organizational interpretations of engine cycling for the $M^{L,R} = 1$ case.

D. The heat flow interpretation

Entropy change always accompanies the transfer of excitations. It is therefore possible to assign an entropy flow to various particle currents, which then takes on the interpretation of heat in classical thermodynamics. To show that this simplified model is appropriately interpreted as an engine/reservoir system, it is necessary to show that the heat flows obey those classically associated with engine cycles.

Because the model has been constructed with a topological correspondence to the thermoacoustic case, there is a natural phenomenology of thermoacoustic engine transport to map to it. The traveling-wave states $+$ and $-$ correspond to the transporting excitations. Classically, the entropy transported from the R to the L reservoir is proportional to the excess number of $+$ over $-$ excitations, while the transport from L to R is proportional to the $-$ over $+$ excess. Meanwhile, the rate of growth of the $+$ excitation number is proportional to the temperature difference of R over L, while the growth rate of $-$ is proportional to the $L-R$ temperature difference. This model is linear, so there will not be the additional proportionality of $+$ and $-$ growth with the current amplitudes of $+$ and $-$.

Only the total particle flow into reservoirs will be directly constrained by the local interaction Hamiltonians. However, at $M^{L,R} = 1$, $\bar{n}_0^L \to \bar{n}^L$, $\bar{n}_0^R \to \bar{n}^R$, while also $(\bar{n}_0^L - \bar{n}_0^R)_t \approx (\bar{n}_l - \bar{n}_r)_t$. By Eqs. (26) and (27),

$$\frac{d}{dt}\left(\frac{\bar{n}^L - \bar{n}^R}{\bar{n}^L + \bar{n}^R}\right) = \nu\left(\frac{\bar{n}_+ - \bar{n}_-}{\bar{n}_+ + \bar{n}_-}\right) \quad (31)$$

and

$$\frac{d}{dt}\left(\frac{\bar{n}_+ - \bar{n}_-}{\bar{n}_+ + \bar{n}_-}\right) = \nu\left(\frac{\bar{n}^R - \bar{n}^L}{\bar{n}^R + \bar{n}^L}\right). \quad (32)$$

Equation (31) assigns particle transport between the reservoirs to an invariant number $\propto \nu$ per traveling-wave cycle. Since $\bar{n}_+ - \bar{n}_-$, the net current, is also the mean number constraint biasing the marginal distribution in the reservoir away from the equilibrium Gibbs distribution, it is the constraint on which the *order* growth in the engine depends. Equation (32) correctly recovers another aspect of the real thermoacoustic phenomenology: that the growth of mean energy in the coherent state is proportional to the driving temperature difference, as for an engine with fixed entropy transport per cycle it must be.

To infer local particle flows between the engine and the reservoir, it is natural to split the interaction Hamiltonian as

$$H_{\mathrm{int}} \equiv H_{\mathrm{int}}^L + H_{\mathrm{int}}^R \quad (33)$$

with

$$H_{\text{int}}^L \equiv -\frac{g}{\sqrt{2}\cos\theta_0}[a_{\theta_0}^\dagger a_0^L + \text{H.c.}] \tag{34}$$

and

$$H_{\text{int}}^R \equiv -\frac{g}{\sqrt{2}\cos\theta_0}[a_{-\theta_0}^\dagger a_0^R + \text{H.c.}]. \tag{35}$$

In the Heisenberg picture, this leads to the expression for the change in the number of particles n^L, due to interaction with n_+, as

$$\frac{d\hat{n}_+^L}{dt} \equiv i[H_{\text{int}}^L, \hat{n}_+], \tag{36}$$

and the change in n^R from interaction with n_+ as

$$\frac{d\hat{n}_+^R}{dt} \equiv i[H_{\text{int}}^R, \hat{n}_+]. \tag{37}$$

The sum of operators interacting with n_+ satisfies the engine conservation law

$$\frac{d\hat{n}_+^L}{dt} + \frac{d\hat{n}_+^R}{dt} = i[H, \hat{n}_+] = -\frac{d\hat{n}_+}{dt}, \tag{38}$$

while the sum for n^L satisfies the reservoir conservation law

$$\frac{d\hat{n}_+^L}{dt} + \frac{d\hat{n}_-^L}{dt} = \frac{d\hat{n}^L}{dt}. \tag{39}$$

Both conservation laws remain true at general $M^{L,R}$, and there are symmetric constructions for the $(-)$ and R sectors, respectively.

Because the entropy change for a thermal distribution is a function only of the mean excitation number, the splitting of particle currents allows a similar splitting of entropy changes into "flows." In a large-$\bar{n}$ limit, where

$$\frac{dS(\bar{n})}{d\bar{n}} \rightarrow \frac{1}{\bar{n}}, \tag{40}$$

the change in the L reservoir entropy from $+$ currents is

$$\frac{dS_+^L}{dt} \equiv \frac{\langle i[H_{\text{int}}^L, \hat{n}_+]\rangle}{\bar{n}^L} \tag{41}$$

and similarly for the R reservoir entropy

$$\frac{dS_+^R}{dt} \equiv \frac{\langle i[H_{\text{int}}^R, \hat{n}_+]\rangle}{\bar{n}^R}. \tag{42}$$

By construction,

$$\frac{dS_+^L}{dt} + \frac{dS_-^L}{dt} = \frac{dS^L}{dt}, \tag{43}$$

and likewise for $L \rightarrow R$.

Since the whole-system sum of marginal entropies is conserved only to second order in fluctuations, it makes sense only to expand Eqs. (41) and (42) to that order, giving

$$\frac{d}{dt}(S_+^L + S_+^R) = -\frac{2}{(\bar{n}^L + \bar{n}^R)}\frac{d\bar{n}_+}{dt} - 2\frac{(\bar{n}^L - \bar{n}^R)}{(\bar{n}^L + \bar{n}^R)^2}$$
$$\times \langle i[(H_{\text{int}}^L - H_{\text{int}}^R), \hat{n}_+]\rangle. \tag{44}$$

In the $g/\sqrt{2}\cos\theta_0 \rightarrow 1$ limit that is dominated by a single mode in both x and y, the operator algebra of Eq. (33) gives the simplified relation

$$i\langle[(H_{\text{int}}^L - H_{\text{int}}^R), (\hat{n}_+ + \hat{n}_-)]\rangle = \nu(\bar{n}_+ - \bar{n}_-), \tag{45}$$

plus error terms of order $1 - g/\sqrt{2}\cos\theta_0$ relative to the terms that are kept. A similar expansion for the antisymmetric number sum gives

$$i\langle[(H_{\text{int}}^L - H_{\text{int}}^R), (\hat{n}_+ - \hat{n}_-)]\rangle$$
$$= 2g[(\bar{n}_y + 1/2) - \tan\theta_0(\bar{n}_x + 1/2)]. \tag{46}$$

Using Eq. (40) for $\bar{n}_+$, it follows that

$$\frac{d}{dt}(S_+^L + S_+^R) = -\frac{dS_+}{dt} + 2g\left(\frac{\bar{n}^L - \bar{n}^R}{\bar{n}^L + \bar{n}^R}\right)$$
$$\times \frac{\tan\theta_0(\bar{n}_x + 1/2) - (\bar{n}_y + 1/2)}{\bar{n}_x + \bar{n}_y}. \tag{47}$$

Classical Carnot's theorem would have Eq. (47) identically zero, because for a reversible process the entropy transport out of R by $(+)$ would exactly equal that into L, so the sum of two *inward* entropy transports would have to vanish. In this problem, Eq. (47) has both symmetric and antisymmetric nonzero terms. The symmetric term is precisely the information (negative of the change in entropy) stored in the $+$ standing wave, the necessary correction to the classical theorem needed to treat the current itself as a thermal object. The second term is totally $\pm$ antisymmetric, and describes an artificial entropy transport which exactly cancels between the two traveling waves, and thus is never actually delivered to either reservoir, or to the engine modes either.

At this point it is useful to make a formal distinction between those entropies considered informational in origin, and those conventionally regarded as thermal. The thermal entropies that pass through an engine come from those correlations necessary to specify the state of the environment. They are originally projected onto the measurements of one reservoir and later transferred to the other's, but never resolved in the engine at any point. The informational entropies are the uncertainties of the actual state of the engine,

which may change if the engine's composition is measured as part of the system characterization.

The informational entropy change is the sum $-dS_+/dt$ $-dS_-/dt$, from Eq. (44) and its $(-)$ counterpart. By Eq. (27), it is quadratic in fluctuations, and it has natural interpretations in terms both of self-organization and of flows. The reduction in $S_+ + S_-$ is just the knowledge gained about the system from the constraint on its current. A state with zero mean current maximizes what would normally be computed as the "equilibrium" free energy. The imposition of a mean current as a constraint would produce just the distribution and entropy computed here as its maximum-ignorance solution. In this case, a constraint on the $\pm$ current excess arises through the dynamics, as the expression of the initial reservoir heterogeneity, and so appears self-generated from the perspective of the time-local coarse graining.

Meanwhile, the thermal entropy passing through the engine is (one half of) the antisymmetric combination

$$\frac{d}{dt}(S_+^L - S_+^R) = \frac{2}{(\bar{n}^L + \bar{n}^R)}\langle i[(H_{\text{int}}^L - H_{\text{int}}^R), \hat{n}_+]\rangle$$
$$+ 2\frac{(\bar{n}^L - \bar{n}^R)}{(\bar{n}^L + \bar{n}^R)^2}\frac{d\bar{n}_+}{dt}. \quad (48)$$

Making use of the same operator identities and single-mode limits as above, this exchanged entropy evaluates to

$$\frac{d}{dt}(S_+^L - S_+^R) = -2g\frac{\tan\theta_0(\bar{n}_x + 1/2) - (\bar{n}_y + 1/2)}{\bar{n}_x + \bar{n}_y}$$
$$+ \nu\left(\frac{\bar{n}_+ - \bar{n}_-}{\bar{n}_+ + \bar{n}_-}\right) - \nu\left(\frac{\bar{n}^L - \bar{n}^R}{\bar{n}^L + \bar{n}^R}\right)^2. \quad (49)$$

The first term is again a $\pm$-antisymmetric combination associated with nonuniform coupling to the standing waves. It could be set to zero with suitable populations of x and y, but never actually accumulates anywhere and is essentially an artifact. The leading-order entropy actually exchanged is linear in $\bar{n}_+ - \bar{n}_-$, thus obeying the thermoacoustic Carnot relation. The quadratic correction is of the same order as the informational entropy, and describes how it is drawn differentially from the two reservoirs.

This model, then, does strictly what was described above. When an appropriate basis is specified, the engine traveling modes appear to transport information about the environment from one reservoir to the other in the process of supporting particle currents. When the temperatures differ, it is not energy, but thermal entropy whose flux is conserved at leading order. The engine uses the resulting excess of energy to augment its own structure, or vice versa. When engine order is growing, though, classical Carnot's theorem is not exactly respected. The entropy the engine rejects at the low-temperature reservoir is slightly greater than what it takes in at high temperature, by just the change in its own structural information.
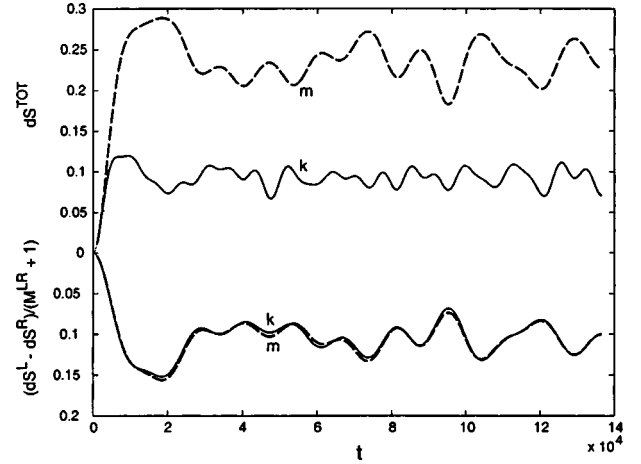

FIG. 4. Changes in total entropy from their initial values (dS), in the combined engine and reservoirs (upper curves), and in transported entropy per degree of freedom $(dS^L - dS^R)/(M^{L,R} + 1)$ (lower curves), versus time in units of cycle periods of the free oscillator. Model values are $M^{L,R} = 30$, $\gamma = 0.3$, $g = 0.1$, and $\tan\theta_0 = 0.995$ or $\nu = 5 \times 10^{-4}$. Initial conditions are $T_0^L = 1100$, $T_0^R = 900$, in energy units where the free oscillator energy is 1. Solid curves use sums of thermal entropies for standing-wave occupation number (k) states in the reservoirs, while dashed curves use position-occupation states m. The $\pm$ basis, which is well behaved at $M^{L,R} = 1$, is used in the engine in all cases. The whole system entropy depends strongly on the representation of the reservoirs; in the k basis, there is no entropy change in the transverse standing waves, which do not project on the $m = 0$ coupling, while the m basis is repopulated dynamically at every position, and thus shows larger overall entropy increase. The *transported* entropy is independent of this distinction to the accuracy of the simulation, and may be regarded as a property of the state of the engine. Despite the complex, apparently irreversible behavior of all of these entropy curves, the temperatures corresponding to the k occupation numbers oscillate harmonically between T_0^L and T_0^R, for all time.

E. More modes and nonlocal state variables

When $M^{L,R} > 1$ or $g/\sqrt{2}\cos\theta_0 \ll 1$, there is no way to induce even quadratic-order entropy conservation, for any sum of marginal entropies computed only from time-local mean excitation numbers. While the sum of reservoir entropies is constant at linear order in fluctuations, the "exchanged" thermal entropy from each reservoir component is proportional to a separate traveling current in the engine. These currents all have different effective temperatures at any time t, even though their distributions combine to form a thermal distribution for total traveling current, which may have yet another, arbitrarily related, temperature. The resulting picture is one in which both classical Carnot's theorem and all informational entropies are lost against a background of entropy fluctuation created by inadequacies of the coarse graining. At early times, the entropy change is always an increase, recovering the usual picture of an irreversible process, as shown in Fig. 4.

This case illustrates the distinction between reversibility defined by conserved indexing of microhistories, as in Sec. II, and defined by the limitations of the experimenter. Under the former, the system is identically reversible, while under

the latter, it would be called irreversible. Yet if one were to Fourier transform the time series of the component occupation numbers, these could be decomposed into the slowly changing occupation numbers of the k states, which together are sufficient state variables for the system at general $M^{L,R}$. Further, while such a decomposition cannot be made *locally* in time, it can be performed within a finite time interval for any fixed $M^{L,R}$. Thus a more sophisticated experimenter [42] would recognize the actual reversibility of the ensemble, while a less sophisticated one would not.

The late-time behavior of Fig. 4 illustrates well the limitations of phenomenological treatment of entropies. There is no violation of the second law implied by the significant decreases of the coarse-grained entropies, since these are loose constraints on the true uncertainty about the distribution. Conversely, a treatment that *defined* entropy functions solely to be spatially and temporally local, and monotonically increasing (as in Ref. [8], Secs. 5.6 and 5.7), to satisfy the second law as if it were a statute, would be a measure of something else besides uncertainty, as entropies are properly defined to be [35].

V. GENERALIZING THE TIME-INDEPENDENT ENTROPY

The high-temperature limit of the linear oscillator model has been explicitly constructed with a partition function indistinguishable from that for an equilibrium system, except in the interpretation of its parameters as temperatures (no longer having only one value) or occupation numbers (having explicit dynamics). The logarithm of the partition function must then yield an effective potential generalizing the Helmholtz equilibrium free energy to one for a microcanonical ensemble of *histories*, in which the constraints from heterogeneity are explicit as in Sec. II.

At this point, it is convenient to introduce some case-specific notation, and to correct a technical omission that has been committed up to now, in computing entropies from the projections of occupation numbers onto engine or reservoir components alone, which differ by factors of 2 from the true occupation numbers of the eigenstates responsible for the oscillation. At $g/\sqrt{2}\cos\theta_0 \to 1$, the sum of entropies computed in this way differs from the true system entropy by constant offsets $\sim \ln 2$ in the high-temperature limit, so the conditions for reversibility and results about entropy transport hold in either variables. However, in order to use the real eigenvalues of the mean-number matrix, it will be assumed that the experimenter has multiplied the component occupations by 2, and performed the trivial integration over the unpopulated state amplitudes to obtain a number matrix with only nonzero eigenvalues. The trace of this projected matrix, which gives the total occupation number, is then denoted $\bar{n}_T \equiv \bar{n}_x + \bar{n}_y$, while the time-independent difference $\bar{n}_{xy} \equiv \bar{n}_x - \bar{n}_y$.

The differences, between l and r or $+$ and $-$ sectors, are functions of time, and as such are not suitable order parameters if not so referenced. Therefore let $\bar{n}_{lr,0} \equiv \bar{n}_l(0) - \bar{n}_r(0)$, denote either the value of the number difference at a reference time $t=0$, or equivalently the whole *function* for

which that difference is the sole constraint. Similarly, let $\bar{n}_{+-,0} \equiv \bar{n}_+(0) - \bar{n}_-(0)$ denote an initial value constraint for $(\pm)$, or the whole history that follows from that constraint.

In terms of these, the projection of the matrix $\bar{n}$ onto the eigenmodes with nonzero population, in an xy basis for the raising and lowering operators, becomes

$$\bar{n}_0 = \frac{1}{2}\begin{bmatrix} \bar{n}_T + \bar{n}_{xy} & \bar{n}_{lr,0} + i\bar{n}_{+-,0} \\ \bar{n}_{lr,0} - i\bar{n}_{+-,0} & \bar{n}_T + \bar{n}_{xy} \end{bmatrix}. \qquad (50)$$

Since these occupation numbers are already true eigenvalues, even though they are referenced to projections onto engine states, they account for all system correlations, so reservoir occupation numbers need not be listed again.

The entropy of the most general configuration constrained only by number density, and represented with a Gaussian-coherent ensemble, is given by Eq. (C12) in terms of $\bar{n}$ of Eq. (50). If an average $\bar{\beta}$ value is defined as

$$\bar{\beta} \equiv \frac{\beta^L + \beta^R}{2}, \qquad (51)$$

it is possible to write the equivalent of an equilibrium free energy for the heterogeneously constrained system as

$$\bar{\beta}\mathcal{A} \equiv \beta^L\bar{n}^L + \beta^R\bar{n}^R - S. \qquad (52)$$

Elementary algebra shows that S is maximized, and $\mathcal{A}$ minimized, at $\bar{n}_{+-,0}=0$. Along this curve, the trace expression (C12) for S factors, and the free energy decomposes to the sum of equilibrium forms

$$\bar{\beta}\mathcal{A} \to \beta^L A_0^L + \beta^R A_0^R. \qquad (53)$$

In this sense, along a hyperplane of the possible classical configurations, the free energies look exactly like their equilibrium counterparts, even though the system is dynamical after $t=0$.

$\mathcal{A}$ is minimized by an order parameter that refers to a time-dependent classical history. It is instructive to ask what would be the closest approximation to this solution, obtained if one were to exclude explicit time dependence from the free energy. In this problem, that amounts to taking some static projection of $\mathcal{A}$, and trying to interpret it as an equilibrium free energy. The value $\bar{\beta}$ of Eq. (51) was chosen so that only the heterogeneous Legendre transform pair, and the dependence of the entropy on the time-varying order parameters, would be excluded from the static projection.

This choice of "effective" equilibrium temperature leads to the expansion

$$\mathcal{A} = 2\bar{n}_T - \frac{1}{\bar{\beta}}S_{\text{stat}} + \left(\frac{\beta^L - \beta^R}{\beta^L + \beta^R}\right)2\bar{n}_{lr,0} - \frac{1}{\bar{\beta}}[S - S_{\text{stat}}],$$
$$(54)$$

where $S_{\text{stat}} \equiv S(\bar{n}_T, \bar{n}_{xy}, 0, 0)$. Equation (54) is the specific instance of Eq. (2) for the linear-oscillator model at $M^{L,R}=1$.

Now, the correct (dynamical) solution from Eq. (53) leads to a total number of particles in x and y of

$$\bar{n}_T \rightarrow \frac{1}{\beta^L} + \frac{1}{\beta^R}, \qquad (55)$$

evenly divided between the two modes and time independent. In contrast, direct minimization of the first two terms of Eq. (54) gives the expected total number

$$\bar{n}_T \rightarrow \frac{2}{\bar{\beta}} = \left(\frac{1}{\beta^L} + \frac{1}{\beta^R} \right) \left[1 - \left(\frac{\beta^L - \beta^R}{\beta^L + \beta^R} \right)^2 \right]. \qquad (56)$$

The result (56) is clearly inconsistent with the x and y particle numbers that would be observed, indicating that the static parts of Eq. (54) are not the "best" equilibrium approximation.

A better approximation could be obtained by setting $2/\bar{\beta} = 1/\beta^L + 1/\beta^R$ and incorporating a constant term $\sim \bar{n}_T(\beta^L - \beta^R)$ into the "dynamic" corrections, so that the static part alone would give the correct value of both x and y numbers from Eq. (56). In that case, however, the l, r, $+$ and $-$ numbers would be equal, and the entropy higher than any of the best-characterized instantaneous entropies in the dynamical case.

This alternative between two mistakes is characteristic of the equilibrium interpretation of systems whose order is maintained by heterogeneous boundary conditions and induced currents. For the naive mean intensive variable, the expected extensive variables are wrongly predicted. For the intensive variable that predicts the correct extensive values, the predicted entropy is higher both than the true value (here the matrix trace), and than the function computed from the measured values of the equilibrium state variables. Section VI will mention some important cases of such misanalysis.

Time- and temperature-dependent potentials

Equation (54) may be evaluated as a generalized Helmholtz free energy of configurations at any time, because the arguments $\bar{n}_{lr,0}$, $\bar{n}_{+-,0}$ refer to complete histories, whose projection onto all such configurations are known from Eq. (22). In a less trivial system, where the evolution of the classical state variables could not be so simply derived, it would be desirable to have an extremization principle that determined the equation of motion, or equivalently the transformation rule for the effective potential.

Such a dynamical principle is derived in Sec. 1 of Appendix F, from the partition function of the preparation, and the time-loop S matrix that defines the microscopic evolution of the eigenstates. A hybrid effective potential is found, whose real part is a generalized Helmholtz potential for the initial conditions, and whose imaginary part is a background-field effective action for the evolution forward from them. The effective potential is slightly different from Eq. (54), being a Legendre transform in the exponentials of the inverse temperatures, rather than in the inverse temperatures themselves, which is conveniently performed on the Gaussian-coherent

representation of the partition function. The initial conditions it specifies are all the same, though. The construction of Sec. 1 of Appendix F gives a general relation between the uncertainty measures over histories examined here, and the classical action from which one generally expects to derive reversible dynamics.

VI. DISCUSSION

A. Relation of entropies to complexity

This paper treats the formation of *order* in reversible systems and, when a component-projected set of state variables can be found, uses the decrease in entropy in some component as the *measure* of its increase in order. In cases like the thermoacoustic engines, where the order accompanies the selection of a phase for the classical wave, creation of order accompanies *innovation* (emergence of the phase as a state variable), and with that *complexity*. It is important to note, though, that order as measured by a decrease in entropy does not necessarily imply increase in complexity [18]. Once a phase emerges as a defined property of the ordered system, further increase of the amplitude of the condensed wave only increases order extensively in the energy, without any accompanying process of emergence that could be interpreted as innovation.

B. Biochemistry and ergodicity

Many treatments of entropy in SO have been motivated by Schrödinger's book [36] which proposes that living systems reject entropy to an environment, in apparent violation of the second law (since the living interior is already more chemically organized), and that this is what distinguishes living from nonliving matter. Analysis of SO in terms of Carnot transport shows clearly how this conclusion can be unwarranted. Both the environment (the reservoirs) and the system (the resonator modes), are themselves *composite* systems. They can sustain either charge or current separation precisely because, uncoupled, there is no conduit for relaxation of either without the other. This kind of coupling is exactly the condition of the cellular interior and its chemical environment, with the cell (or mitochondrial) membrane mediating much of the energy transfer. The dynamics of engines can reject entropy to the environment while lowering free energy, but not relative to those states that the system and environment could attain independently.

Said more generally, while the *joint* states of the system (the cellular interior) and the environment may evolve ergodically, thus permitting the substitution of time for ensemble averages, the *component* projections of those states do not evolve ergodically independently of each other. Thus it is important to separate the intrinsic uncertainty of an ensemble (such as the set of viable molecular arrangements in a particular cell), from ergodicity of the dynamics, which may be very different for parts than for the whole.

Another feature of the model used in this paper was specifically chosen to match the energetic structure of photosynthetic life. Without biochemistry, the electronic transitions in terrestrial matter excitable by visible light are largely un-

coupled by quantum selection rules from the vibrational and rotational excitations of the 300 K ambient microwave spectrum. This state corresponds to the preparation conditions in the model, where orthogonal engine states (the $\pm\theta_0$ standing waves) are, respectively, in equilibrium with their reservoirs at different temperatures, but induce no coupling between the reservoirs. The emergence of biochemistry solves the difficult problem of overcoming quantum selection rules with high throughput, converting visible photons to redox couples [37], whose relaxation in steady state ultimately reradiates entirely in the vibrational and rotational bands.

C. Generalized flow ground states

It has been demonstrated in this paper that currents can be classical state variables, and that, in general, they are the conjugate extensive variables to heterogeneous temperatures or other thermal potentials imposed by the environment. It is natural to call the extrema of these dynamical effective potentials *generalized flow ground states*. They have the same relation to microhistories as ground states have to configurations in equilibrium, namely, as minimal sufficient statistics [17]. When they can be computed, they have the important feature of *quantifying* the energetic preference for states as a function of flow variables, as it has so usefully been computed for states in equilibrium.

ACKNOWLEDGMENTS

APPENDIX A: HARMONIC OSCILLATOR DEFINITIONS AND NOTATION

This appendix establishes definitions and notation for one-dimensional (scalar) and M-dimensional (vector) simple quantum harmonic oscillators. Special attention is given to coherent states, and convenient ways of constructing them in the vector case.

1. One-dimensional oscillation

The algebra of the linear, one-dimensional quantum harmonic oscillator is generated by a raising operator $a^{\dagger}$, and its Hermitian conjugate lowering operator, a [38]. These are normalized by the commutation relation

$$[a,a^{\dagger}]=1. \tag{A1}$$

In terms of these the position operator is

$$\hat{x}\equiv\tfrac{1}{2}(a+a^{\dagger}), \tag{A2}$$

and the momentum operator

$$\hat{p}\equiv\frac{i}{2}(a-a^{\dagger}). \tag{A3}$$

The number operator is defined as

$$\hat{n}\equiv a^{\dagger}a. \tag{A4}$$

The interpretation of $a^{\dagger},a$ as creation and annihilation operators [24] will be natural in describing the effective theory of engines or LC circuits, in which case $\hat{x}$ and $\hat{p}$ become the field operator and its conjugate momentum.

The one-particle vacuum state is defined by the operation $a|0\rangle\equiv0|0\rangle$, and the eigenstates of the number operator are

$$|n\rangle\equiv\frac{(a^{\dagger})^n}{\sqrt{n!}}|0\rangle. \tag{A5}$$

The commutation relation (A1) then gives

$$\hat{n}|n\rangle\equiv n|n\rangle. \tag{A6}$$

A scalar *coherent state* is defined for any complex number ξ as

$$|\xi\rangle\equiv e^{-|\xi|^2/2}\sum_{n=0}^{\infty}\frac{\xi^n(a^{\dagger})^n}{n!}|0\rangle\equiv e^{-|\xi|^2/2}\sum_{n=0}^{\infty}\frac{\xi^n}{\sqrt{n!}}|n\rangle. \tag{A7}$$

It is an eigenstate of the annihilation operator, with

$$a|\xi\rangle\equiv\xi|\xi\rangle. \tag{A8}$$

From the definitions of the position and momentum operators, it follows that the expectations

$$\langle\xi|\hat{x}|\xi\rangle=\mathrm{Re}(\xi) \tag{A9}$$

and

$$\langle\xi|\hat{p}|\xi\rangle=\mathrm{Im}(\xi). \tag{A10}$$

The coherent states correspond to "classical" wave packets under the correspondence principle, and the expectation values (A9) and (A10) evolve according to the equations of motion of a classical simple harmonic oscillator.

The coherent-state expectation of the quantized version of a classical Hamiltonian is

$$\langle\xi|\hat{x}^2+\hat{p}^2|\xi\rangle=\langle\xi|\hat{n}|\xi\rangle+\tfrac{1}{2}=|\xi|^2+\tfrac{1}{2}. \tag{A11}$$

The expected excitation number is $|\xi|^2$. Excitation number in coherent states is Poisson distributed, with the probability of number n in state ξ defined and evaluated as

$$P_{\xi}(n)\equiv|\langle\xi|n\rangle|^2=e^{-|\xi|^2}\frac{|\xi|^{2n}}{n!}. \tag{A12}$$

2. Oscillation in more than one dimension

Vector simple harmonic oscillators are defined by raising and lowering operators with subscript m, which create excitations independently in some number M of dimensions. The vector ground state will be denoted as before, and eigenstates of the vector number operator are now indexed by a vector-

valued excitation number, which will again be denoted n. With proper contraction rules of row with column vectors, this will create no confusion, because the vector oscillator is an exact formal extension of the scalar case.

Vector number states are created by the product of raising operators

$$|n\rangle \equiv \prod_{m=1}^{M} \frac{(a^{\dagger})^{n_m}}{\sqrt{n_m!}}|0\rangle. \tag{A13}$$

Coherent states are indexed by a vector ξ with complex coefficients, and created from the vacuum by

$$|\xi\rangle \equiv \prod_{m=1}^{M} e^{-|\xi^m|^2/2} \sum_{n_m=0}^{\infty} \frac{(a_m^{\dagger}\xi^m)^{n_m}}{n_m!}|0\rangle. \tag{A14}$$

The basis of excitations indexed by m is called *orthogonal* if raising and lowering operators at different m commute. In an orthogonal basis, it is very convenient to define the contraction $a_\xi^{\dagger} \equiv a_m^{\dagger}\xi^m$ as the raising operator for an excitation along direction ξ. These may or may not be canonically normalized, depending on the value of $|\xi|^2$. The reason to introduce them is that the multinomial expansion

$$(a_\xi^{\dagger})^N = \sum_{\{n;\Sigma_{m=1}^{M} n_m = N\}} \frac{N!}{n_1!, \ldots, n_M!} \prod_{m=1}^{M} (a_m^{\dagger}\xi^m)^{n_m} \tag{A15}$$

implies the compact representation

$$|\xi\rangle \equiv e^{-|\xi|^2/2} \sum_{N=0}^{\infty} \frac{(a_\xi^{\dagger})^N}{N!}|0\rangle. \tag{A16}$$

Thus any vector coherent state may be regarded as a scalar coherent state created by the appropriate raising operator. The magnitude of ξ occurring in the normalization is just the scalar product $|\xi|^2 \equiv \xi^{\dagger}\xi$.

3. Basis transformation and time evolution

The full commutation algebra in an orthogonal basis of canonically normalized raising and lower operators is defined to be

$$[a^m, a_n^{\dagger}] = \delta_n^m. \tag{A17}$$

The lowering operators may be transformed to any other basis by a unitary transformation $a^{\mu} \equiv v^{\mu}_{\ m}a^m$, if the corresponding raising operators undergo the inverse transformation $a_\nu^{\dagger} \equiv a_n^{\dagger}v^n_{\ \nu}$. It then follows that in the new basis

$$[a^{\mu}, a_\nu^{\dagger}] = v^{\mu}_{\ m}[a^m, a_n^{\dagger}]v^n_{\ \nu} = v^{\mu}_{\ m}v^m_{\ \nu} = \delta_\nu^{\mu}, \tag{A18}$$

so the transformed operators are again orthogonal and canonically normalized. In the appendixes and the text, superscript greek indices will be reserved for operators which create eigenstates of some Hamiltonian, and subscript roman indices will denote all other bases. Typically these will be bases in which a coupled system factors into engine and

reservoir components, and in which heterogeneous preparation conditions are block diagonal. Geometric objects like $a_\xi^{\dagger}$ have component representations in any basis, and thus define the conjugate transformation rules for the complex vector ξ: $a_\xi^{\dagger} = a_n^{\dagger}v^n_{\ \mu}v^{\mu}_{\ m}\xi^m \equiv a_\mu^{\dagger}\xi^{\mu}$.

The number operator for excitations of Hamiltonian eigenstate μ is defined as $\hat{n}_\mu \equiv a_\mu^{\dagger}a^{\mu}$ (no sum). The Hamiltonian assigns energy to μ excitations as

$$[H, a_\mu^{\dagger}] \equiv E_\mu a_\mu^{\dagger}, \tag{A19}$$

and hence can be written

$$H = \sum_{\mu} E_\mu \hat{n}_\mu. \tag{A20}$$

States evolve in the Schrödinger picture under the time-evolution operator e^{iHt}. When this is applied to coherent states, the subscript t will be introduced as the time index, so that

$$|\xi_t\rangle \equiv e^{iHt}|\xi_0\rangle. \tag{A21}$$

$|\xi_t\rangle$ is created at any time by the same relation (A16), with time evolution introducing only the phase shifts

$$\xi_t^{\mu} \equiv e^{iE_\mu t}\xi_0^{\mu}. \tag{A22}$$

APPENDIX B: DENSITY MATRICES AND COARSE GRAININGS

It is useful to introduce the definitions of coarse graining, and the examples that will be used in the text, because probability notation arises that will be used in later appendixes. The starting definition is that, for $\{|\psi\rangle\}$, some collection of quantum states, any density matrix can be written as a sum of outer products

$$\rho \equiv \sum_{\psi} \rho_\psi |\psi\rangle\langle\psi|. \tag{B1}$$

A *coarse graining* of ρ is a map from ρ to some other density $\tilde{\rho}$ which averages out some of the information in ρ [14]. A particular map used in the text will be called the *annular* coarse graining, defined in terms of a set of number states $|n\rangle$ by

$$\tilde{\rho}_A \equiv \sum_{n} \text{Tr}(\rho|n\rangle\langle n|)|n\rangle\langle n| \equiv \sum_{n} P_\rho(n)|n\rangle\langle n|. \tag{B2}$$

This map removes information in the relative phases of different $|n\rangle$ components, in which ρ may not be diagonal. The nature of the averaging can be understood by applying it to the outer product of a coherent state. Such a product corresponds to a ball in a classical phase space, the x and p values of whose center are the real and imaginary parts of some complex vector ξ. If this ball represents ρ, the coarse-grained density $\tilde{\rho}_A$ uniformly populates the annulus in the

phase space with mean radius $|\xi|$, and radial variance comparable to that in the original ρ.

A few lines of algebra show that map (B2) manifestly satisfies the two conditions on a coarse graining set forth in Ref. [14]. It is idempotent,

$$\widetilde{\widetilde{\rho}} = \widetilde{\rho}, \qquad (B3)$$

and states made typical by the coarse-grained distribution are also typical in the original (fine-grained) distribution:

$$\mathrm{Tr}(\rho \ln \widetilde{\rho}) = \mathrm{Tr}(\widetilde{\rho} \ln \widetilde{\rho}). \qquad (B4)$$

The *entropy* of any density is defined as

$$S_\rho \equiv -\mathrm{Tr}(\rho \ln \rho), \qquad (B5)$$

so it follows that the entropy of the coarse-grained density under the same definition is

$$S_{\widetilde{\rho}_A} \equiv -\sum_n P_\rho(n) \ln P_\rho(n), \qquad (B6)$$

a function only of the occupation-number probabilities. The conditions defining a coarse graining, together with the concavity of the logarithm, imply that $S_{\widetilde{\rho}_A} \geq S_\rho$, for any ρ.

A second stage of coarse graining can be applied to ρ, by *marginalization* of $\widetilde{\rho}_A$. The marginal probability of occupation n_m of some single component m in the density ρ is defined as

$$\widetilde{P}_\rho(n_m) \equiv \sum_{n_{k \neq m}} P_\rho(n). \qquad (B7)$$

Marginalization of the probability of vector excitation number n is replacement of the joint probability with the product of component marginals, denoted

$$\widetilde{P}_\rho(n) \equiv \prod_m \widetilde{P}_\rho(n_m). \qquad (B8)$$

The marginal coarse graining of ρ is defined as

$$\widetilde{\rho}_M \equiv \sum_n \widetilde{P}_\rho(n) |n\rangle\langle n|. \qquad (B9)$$

Marginalization produces an entropy that is a sum of the marginal entropies of each component m:

$$S_{\widetilde{\rho}_M} \equiv \sum_m \left[-\sum_{n_m} \widetilde{P}_\rho(n_m) \ln \widetilde{P}_\rho(n_m) \right]. \qquad (B10)$$

This coarse graining is performed whenever an interacting engine/reservoir system is factored into separate "engine" and "reservoir" components, which are assumed to have independently well-defined entropies and equations of state. The marginalization need not be complete, though when m denotes intracomponent eigenstates that are coupled only through the inter-component interactions, it effectively is.

APPENDIX C: THERMAL DENSITIES AND THEIR COHERENT-STATE REPRESENTATIONS

All the classical states in this paper are built from thermal density matrices. These are maximum-ignorance distributions consistent with fixed expected energy [14,19], and are defined in terms of an inverse temperature β as

$$\rho_\beta \equiv \frac{1}{Z} \sum_n |n\rangle\langle n| \exp\left(-\beta \sum_\mu n_\mu E_\mu \right). \qquad (C1)$$

The normalization factor Z is called the *partition function*, and equals

$$Z = \sum_n \exp\left(-\beta \sum_\mu n_\mu E_\mu \right). \qquad (C2)$$

Since the energy eigenvalues E_μ will be given units of frequency, β will have units of time.

Thermal densities may alternatively be written as integrals over outer products of coherent states. In a basis of eigenstate excitations, the vector thermal density takes the form

$$\rho_\beta = \frac{1}{Z} \int \left(\prod_\mu \frac{e^{\beta E_\mu}}{\pi} d\xi_\mu^* d\xi^\mu \right) e^{-\xi_\mu^* K_\nu^\mu \xi^\nu} |\xi\rangle\langle\xi|. \qquad (C3)$$

A kernel matrix K is introduced by the Gaussian integral, which is diagonal in the eigenstate basis, with eigenvalues $K_\nu^\mu \equiv \delta_\nu^\mu (e^{\beta E_\mu} - 1)$. This K may be checked to recover the thermal occupation number probabilities, by evaluating the trace defined in Eq. (B2),

$$\widetilde{P}_{\rho_\beta}(n_\mu) = \frac{1}{Z_\mu n_\mu!} \int_0^\infty e^{\beta E_\mu} d|\xi^\mu|^2 \exp(-e^{\beta E_\mu} |\xi^\mu|^2) |\xi^\mu|^{2n_\mu}. \qquad (C4)$$

Here all ξ component integrals evaluate to one except at μ, and the normalization Z_μ is the partition function for the density over the μ eigenstate alone. The exponential integral differs from $n_\mu!$ only by the normalization $\exp(-\beta n_\mu E_\mu)$, which with Z_μ recovers the thermal distribution.

The mean excitation number at any μ is similarly easy to evaluate by Gaussian integration, as the trace

$$\mathrm{Tr}(\rho_\beta \hat{n}_\mu) = \frac{1}{Z} \int \left(\prod_\mu \frac{e^{\beta E_\mu}}{\pi} d\xi_\mu^* d\xi^\mu \right) e^{-\xi_\mu^* K_\nu^\mu \xi^\nu} |\xi^\mu|^2$$

$$\equiv \overline{n}_\mu = (K^{-1})_\mu^\mu, \qquad (C5)$$

where a notation $\overline{n}_\mu$ has been introduced for the mean. The population is the inverse of the μ eigenvalue of K, the correct thermal result.

The entropy of a thermal density is a sum of marginal entropies in the eigenstate basis. Evaluating these as functions of the mean occupation numbers gives

$$S_{\rho_\beta} = \sum_\mu S(\overline{n}_\mu) \equiv \sum_\mu (\overline{n}_\mu + 1) \ln(\overline{n}_\mu + 1) - \overline{n}_\mu \ln \overline{n}_\mu. \qquad (C6)$$

The preceding equations reduce the vector problem to a collection of simple scalar evaluations, but the physical content of any interesting problem is much more apparent in a geometric notation. The Gaussian kernel can be written in the basis-independent form

$$\xi_\mu^* K_\nu^\mu \xi^\nu = \xi^\dagger K \xi. \tag{C7}$$

The associated measure over complex vectors ξ is invariant under unitary transformations, and so is also defined from the product measure in the eigenstate basis as

$$\prod_\mu d\xi_\mu^* d\xi^\mu = d\xi^\dagger d\xi. \tag{C8}$$

In this geometric representation, the thermal density (C3) becomes

$$\rho_\beta = \frac{\text{Det}(K+I)}{Z\pi^M} \int d\xi^\dagger d\xi e^{-\xi^\dagger K \xi} |\xi\rangle\langle\xi|, \tag{C9}$$

while the partition function is the ratio of determinants

$$Z = \frac{\text{Det}(K+I)}{\pi^M} \int d\xi^\dagger d\xi e^{-\xi^\dagger K \xi} = \frac{\text{Det}(K+I)}{\text{Det}K}. \tag{C10}$$

The eigenstate occupation numbers are eigenvalues of the diagonal matrix K^{-1}, which thus defines a basis-independent mean number matrix

$$\bar{n} \equiv K^{-1}. \tag{C11}$$

The thermal entropy (C6) then has a representation which is manifestly invariant under unitary transformation of the ξ basis,

$$S_{\rho_\beta} = \text{Tr}[(\bar{n}+1)\ln(\bar{n}+1) - \bar{n}\ln\bar{n}]. \tag{C12}$$

It is immediately apparent that thermal density matrices are a proper subset of the Gaussian-coherent density matrices, and that Eqs. (C9)–(C12) hold for a general kernel matrix K with all positive eigenvalues. In particular, the trace form for the entropy can always be reduced to a sum over marginals of the eigenvalues of K, which need not be those of any Hamiltonian. This property of Gaussian-coherent representations will furnish a very easy way to impose conditions of heterogeneous temperature, on system components whose basis states do not diagonalize the fully interacting Hamiltonian.

APPENDIX D: GAUSSIAN COHERENT REPRESENTATIONS GIVE THERMAL MARGINALS

A property of Gaussian-coherent densities, which will be useful in the analysis of the model engine, is that all of their marginals are exactly thermal, under any basis related to the eigenstates of K by unitary transformation. This will lead to surprising ways of hiding order when the unitary transforma-

tions are generated by the system's own dynamical evolution.

To prove this result, define a basis-independent raising operator $a_\sigma^\dagger \equiv a^\dagger \sigma \equiv a_\mu^\dagger \sigma^\mu$, in terms of an arbitrary complex vector σ normalized to $\sigma^\dagger \sigma = 1$. The conjugate lowering operator is $a_\sigma \equiv \sigma^\dagger a \equiv \sigma_\mu^* a^\mu$, and the number operator for excitations along the σ direction is $\hat{n}_\sigma \equiv a_\sigma^\dagger a_\sigma$. Using representation (A16) for the coherent state $|\xi\rangle$, the mean σ-excitation number in density ρ_β evaluates simply to

$$\text{Tr}(\rho_\beta \hat{n}_\sigma) = \frac{\text{Det}K}{\pi^M} \int d\xi^\dagger d\xi e^{-\xi^\dagger K \xi} |\sigma^\dagger \xi|^2 = \sigma^\dagger K^{-1} \sigma, \tag{D1}$$

the generalization of Eq. (C5).

Meanwhile, the marginal probability of exactly n_σ excitations extracts only n_σ powers of the σ component of ξ, generalizing Eq. (C4). A weight factor is added to the Gaussian kernel from the normalization of the σ component of $|\xi\rangle$, which cancels against the polynomial sum in all other orthogonal components. The resulting expectation is an elementary Gaussian integral generalizing the Gamma function of the scalar case:

$$\tilde{P}_\rho(n_\sigma) = \frac{\text{Det}K}{\pi^M n_\sigma!} \int d\xi^\dagger d\xi e^{-\xi^\dagger K \xi - |\sigma^\dagger \xi|^2} |\sigma^\dagger \xi|^{2n_\sigma}$$

$$= \frac{\text{Det}K}{\text{Det}(K+\sigma\sigma^\dagger)} [\sigma^\dagger (K+\sigma\sigma^\dagger)^{-1}\sigma]^{n_\sigma}. \tag{D2}$$

Its important property is that $\tilde{P}_\rho(n_\sigma)$ is properly normalized, while the ratio at different values of n_σ is a power of $\sigma^\dagger(K+\sigma\sigma^\dagger)^{-1}\sigma$, making the distribution exponential in n_σ, or thermal. It is unnecessary to evaluate the more complex matrix inverse $(K+\sigma\sigma^\dagger)^{-1}$, as it is related to the mean excitation number by

$$\frac{\sigma^\dagger(K+\sigma\sigma^\dagger)^{-1}\sigma}{1+\sigma^\dagger(K+\sigma\sigma^\dagger)^{-1}\sigma} = \sigma^\dagger K^{-1}\sigma. \tag{D3}$$

Similarly, the normalization of the marginal distribution has the simple evaluation

$$\frac{\text{Det}K}{\text{Det}(K+\sigma\sigma^\dagger)} = \frac{1}{1+\sigma^\dagger K^{-1}\sigma}. \tag{D4}$$

APPENDIX E: PRODUCT-THERMAL INITIAL CONDITIONS

Just as all marginal distributions from a Gaussian-coherent density are thermal, arbitrary heterogeneous thermal initial conditions can be imposed with such a density, in any basis related by unitary transformation to the eigenstate basis. For this appendix, suppose that the coordinates ξ define the Gaussian integral for the raising operators of the model in Sec. IV.

In the model it was possible to choose a "preparation

Hamiltonian'' (13) whose eigenstates were products of eigenstates in left and right sectors, denoted L and R. Each sector included all of the excitations in its respective reservoir, and a sector-unique linear combination of the excitations in the engine. With respect to this decomposition, write the column vector of complex ξ coefficients

$$\xi = \begin{bmatrix} \xi^L \\ \xi^R \end{bmatrix}. \qquad \text{(E1)}$$

This basis decomposition is related by an orthogonal transformation to $\{x,S,y,A\}$ basis of eigenstate excitations of the fully interacting engine.

For the subset of coefficients ξ^L, a standard thermal density matrix is given by

$$\rho_{\beta^L}^L = \frac{\text{Det}(K^L + I)}{Z^L \pi^{M^L}} \int d\xi^{L\dagger} d\xi^L e^{-\xi^{L\dagger} K^L \xi^L} |\xi^L\rangle\langle\xi^L|, \quad \text{(E2)}$$

in terms of inverse temperature β^L, as per Eq. (C9). The corresponding density for ξ^R is given in terms of a β^R as

$$\rho_{\beta^R}^R = \frac{\text{Det}(K^R + I)}{Z^R \pi^{M^R}} \int d\xi^{R\dagger} d\xi^R e^{-\xi^{R\dagger} K^R \xi^R} |\xi^R\rangle\langle\xi^R|. \quad \text{(E3)}$$

Coherent states for the full system are products of coherent states for the L and R factors, by application of the Binomial theorem to Eq. (A16). Thus,

$$|\xi\rangle = |\xi^L, \xi^R\rangle, \qquad \text{(E4)}$$

and the product density as long as the two sectors are decoupled is simply

$$\rho = \rho_{\beta^L}^R \rho_{\beta^R}^R. \qquad \text{(E5)}$$

Product (E5) is itself a Gaussian integral over states (E4). If the sector coefficients are reassembled into the column vector (E1), the kernel of that integral has the block-diagonal form

$$K = \begin{bmatrix} K^L & \\ & K^R \end{bmatrix}. \qquad \text{(E6)}$$

As long as the system is evolving under the preparation Hamiltonian, the phases of the separate ξ^L and ξ^R cancel from their respective Gaussian integrals, because K^L and K^R are diagonal in the sector eigenbases. Thus the time index need not be specified explicitly for either the factor or product densities to be well defined.

To use density (E5) to specify the same distribution at other times, suppose that interactions are turned on at some time labeled 0. Then take the coefficient vector ξ to be the explicit coherent-state vector ξ_0 of Eq. (A22). The evolving density specified by these heterogeneous initial conditions is defined at all later times by taking the Gaussian integral over ξ_0, with K fixed, and allowing the coherent states $|\xi\rangle$ to evolve as in the Schrödinger picture. This is how a fixed

uncertainty specifies a distribution on which measurements may be made over an indefinite time interval.

Alternatively, something like the Heisenberg picture may be adopted, by taking the Gaussian integral over the same coefficients ξ_t used to index the coherent states. The preparation basis is related to the eigenbasis by an orthogonal transformation, and the eigenbasis evolves under the diagonal time-evolution matrix. Thus the preparation basis at time 0 is related to the eigenbasis at any other time by a unitary transformation, and the measure is invariant under these transformations. Thus the only change in the Gaussian integral is by similarity transform of the kernel K, from the block-diagonal form (E6) at time 0 to whatever basis is desired at time t. The inverse matrix $\bar{n}_t$ evolves under the identical similarity transform, worked out for the examples of interest in the text in Eq. (22).

APPENDIX F: CONNECTION TO THE CLASSICAL ACTION

Single-time correlations in linear models are elementary to evaluate in Gaussian-coherent ensembles, because the fields ξ may simply be transformed from one time t to another t' with the unitary matrix $e^{iE(t'-t)}$. The measure is invariant, and the kernel evolves by similarity transformation under the same unitary matrix, as noted in Appendix E.

An alternative approach is to explicitly embed the appropriate time-evolution operator (called the time-loop S matrix) in the thermal trace, and compute classical expectations in a background field expansion. While more cumbersome for single-time correlations of linear systems, this method automatically includes the correlations implied by classical dynamics at arbitrary collections of times, and provides a systematic perturbative construction for systems that cannot be solved exactly.

The background field expansion of heterogeneous-temperature systems is of interest as a formal construction in its own right, because it produces a sum of a classical effective action and a multitemperature thermal effective potential. The former identifies classical dynamics, while the latter specifies initial conditions. The sum of the two potentials has a natural interpretation as a stationary-point evaluation of a Matsubara path integral, analytically continued to real times and extended to incorporate multiple temperatures. Both the interpretation and the form are retained even if the system is nonlinear. Thus, even if one does not choose to carry through the full construction in complex cases, if the classical action is known, the justification is provided for combining it with a thermal ensemble to select nontrivial initial conditions, *along with a rule for transforming the combination to describe origins of equivalent dynamics at different times.*

1. Forming the time-loop generating functional

The thermal ensemble is specified at some time $t=0$. Coherent-state parameters are labeled ξ_0, as in Appendix E, and the kernel for the $t=0$ distribution is denoted K_0.

Correlations at arbitrary times are then computed by inserting Schrödinger-picture operators at those times into a

thermal trace of the S matrix that evolves the states at $t=0$ over an interval sufficient to include them. In zero-temperature field theory, the S matrix generally interpolates between the initial and final true ground states (in the many-particle application, "in" and "out" vacua). Since both bra and ket states in the thermal ensemble are given at $t=0$, the "final" ket is defined by the action of the S matrix on the $t=0$ ket, which is the conjugate of the "initial" bra. Since this evolution must be free from the action of the operators to be inserted, there are two S matrices forming a time loop [39], with all physically correlated observables inserted into only one leg.

The time-loop S matrix is the time-loop-ordered product of the sequence of infinitesimal operators $e^{i\hat{H}dt}$, where $\hat{H}$ includes interaction terms if the theory has them. dt is positive on the forward leg, and negative on the return. If we wish to make a background-field theory of the number operators, their expectations can be generated with a perturbation of the energy matrix by an arbitrary probe $J(t)$. The generating functional then has the form

$$\zeta_J \equiv Z \, \mathrm{Tr}\left\{ \rho_\beta \boldsymbol{T} \exp\left(i \oint_0 dt \, \mathrm{Tr}[(E-J)\hat{\boldsymbol{n}}] \right) \right\}, \quad (F1)$$

where the Schrödinger-picture number operator is defined at each time as

$$\hat{\boldsymbol{n}} = [\hat{\boldsymbol{n}}_\mu^\nu] \equiv [a_\mu^\dagger a^\nu], \quad (F2)$$

so that the Hamiltonian is

$$\hat{H} \equiv \mathrm{Tr}[E\hat{\boldsymbol{n}}]. \quad (F3)$$

$\boldsymbol{T}$ denotes time-loop ordering, and the $\oint_0$ indicates that the expansion of the exponential begins and ends at $t=0$. The outer trace in Eq. (F1) is over quantum states, and the trace in the exponential is simply over matrix indices, as in Eq. (F3).

Since coupling to the number operator preserves the linearity of the theory, the distributions of all single-component occupation numbers in the perturbed theory remain exactly thermal. The perturbed S matrix continues to act by matrix multiplication of the parameter ξ, so from form (C9) and algebra, one can write

$$\zeta_J = \frac{\mathrm{Det}(K_0+I)}{\pi^M} \int d\xi_0^\dagger d\xi_0 \exp\left[-\xi_0^\dagger \left(K_0+I \right. \right.$$
$$\left. \left. -\boldsymbol{T} \exp\left(i \oint_0 dt(E-J) \right) \right) \xi_0 \right]. \quad (F4)$$

First variation of ζ_J generates the expectation of $\hat{\boldsymbol{n}}$ in the presence of J, denoted $\bar{n}_J$:

$$i\frac{\delta \ln \zeta_J}{\delta J(t)} = \frac{Z}{\zeta_J}\mathrm{Tr}\left\{ \rho_\beta \boldsymbol{T} \exp\left(i \oint_0 dt \, \mathrm{Tr}[(E-J)\hat{\boldsymbol{n}}] \right) \hat{\boldsymbol{n}}(t) \right\}$$

$$\equiv \bar{n}_J(t)$$

$$= \boldsymbol{T} \exp\left(i \int_0^t dt'(E-J) \right)$$

$$\times \left[K_0+I-\boldsymbol{T}\exp\left(i \oint_0 dt(E-J) \right) \right]^{-1} \boldsymbol{T}$$

$$\times \exp\left(i \int_t^0 dt'(E-J) \right). \quad (F5)$$

Evolution under the free Hamiltonian is simple to factor out of perturbative expansions in J, so it is convenient to compute the so-called "interaction picture" number matrix

$$\tilde{n}_J(t) \equiv e^{-iEt}\bar{n}_J(t)e^{iEt}, \quad (F6)$$

and its conjugate probe matrix

$$\tilde{J}(t) \equiv e^{-iEt}J(t)e^{iEt}. \quad (F7)$$

Computing with interaction-picture operators is equivalent to performing the unitary transformation of the ξ variables to appropriate times at the outset of the calculation. As expected, in the linear theory this will account for all dynamical structure, and the rest of the background field expansion will simply verify the constancy of the interaction-picture observables. In a nontrivial theory, only an exactly solvable part of the Hamiltonian would be used to define the interaction picture, and the remaining interactions would have to be treated perturbatively together with the influence of J.

The time-loop-ordered matrix appearing in Eq. (F4) has the expansion in interaction picture $\tilde{J}$,

$$\boldsymbol{T}\exp\left(i \oint_0 dt(E-J) \right) = \boldsymbol{T}\exp\left(-i \oint_0 dt \, \tilde{J} \right), \quad (F8)$$

while the similar expansion for $\tilde{n}_J$ is

$$\tilde{n}_J(t) = \boldsymbol{T}\exp\left(-i \int_0^t dt' \, \tilde{J} \right)\left[K_0+I-\boldsymbol{T}\exp\left(-i \oint_0 dt \, \tilde{J} \right) \right]^{-1} \boldsymbol{T}$$

$$\times \exp\left(-i \int_t^0 dt' \, \tilde{J} \right). \quad (F9)$$

At $t=0$ (either initial or final), Eq. (F6) is no transformation, so

$$\bar{n}_J(0_-) = \tilde{n}_J(0_-)$$

$$= \boldsymbol{T}\exp\left(-i \oint_0 dt \, \tilde{J} \right)$$

$$\times \left[K_0+I-\boldsymbol{T}\exp\left(-i \oint_0 dt \, \tilde{J} \right) \right]^{-1}, \quad (F10)$$

$$\bar{n}_J(0_+) = \tilde{n}_J(0_+) = \left[K_0 + I - \boldsymbol{T}\exp\left(-i\oint_0 dt\,\tilde{J}\right)\right]^{-1}\boldsymbol{T}$$

$$\times \exp\left(-i\oint_0 dt\tilde{J}\right), \tag{F11}$$

and when $\tilde{J}$ is Hermitian, $\bar{n}_J(0_-)$ and $\bar{n}_J(0_+)$ are related by a unitary transformation. At every time, $\tilde{n}_J$ evolves according to

$$\frac{d\tilde{n}_J(t)}{dt} = -i[\tilde{J}(t), \tilde{n}_J(t)]. \tag{F12}$$

Finally, it is convenient to use Eq. (F8) to evaluate ζ_J in closed form as

$$\zeta_J = \frac{\mathrm{Det}(K_0+I)}{\pi^M}\int d\xi_0^\dagger d\xi_0$$

$$\times \exp\left[-\xi_0^\dagger\left(K_0+I-\boldsymbol{T}\exp\left(-i\oint_0 dt\,\tilde{J}\right)\right)\xi_0\right]$$

$$= \frac{\mathrm{Det}(K_0+I)}{\mathrm{Det}\left[K_0+I-\boldsymbol{T}\exp\left(-i\oint_0 dt\,\tilde{J}\right)\right]}$$

$$= Z\,\mathrm{Det}[K_0\langle\xi_0\xi_0^\dagger\rangle_J]$$

$$= Z\,\mathrm{Det}[K_0\bar{n}_J(0)]\exp\left(i\oint dt\,\mathrm{Tr}(\tilde{J})\right), \tag{F13}$$

where the expectation $\langle\xi_0\xi_0^\dagger\rangle$ is with respect to the Gaussian integral.

2. Subtleties of the Legendre transformation

Ordinarily, i times the effective action is defined as the Legendre transform of the logarithm of the generating functional from the J-perturbed S matrix. In the finite-temperature problem here, though, it is not obvious even whether that transform should be imaginary or real. Sorting out the correct way to do the transformation will expose subtleties in the inversion of the relation n_J that encodes initial conditions as well as dynamics.

The reason there is a puzzle is that Eq. (F13) clearly has the form

$$\zeta_J = \frac{\mathrm{Det}(K_0+I)}{\pi^M}\int d\xi_0^\dagger d\xi_0\, e^{-\xi_0^\dagger(K_0+J)\xi_0} \tag{F14}$$

with the matrix

$$\boldsymbol{J} = I - \boldsymbol{T}\exp\left(-i\oint_0 dt\,\tilde{J}\right). \tag{F15}$$

If $\boldsymbol{J}$ is Hermitian (a nontrivial restriction on J, but one which is useful to impose for a moment in order to see the point),

the generating functional is purely real, and its proper Legendre transform would seem to be

$$\zeta_{\bar{\nu}_0} = e^{\mathrm{Tr}[J\bar{\nu}_0]}\zeta_J|_{J(\bar{\nu}_0)}$$

$$= e^{\mathrm{Tr}[I-K_0\bar{\nu}_0]}\zeta_{J(\bar{\nu}_0)}$$

$$= Z\,e^{-\mathrm{Tr}[K_0\bar{\nu}_0-I]}\mathrm{Det}[K_0\bar{\nu}_0], \tag{F16}$$

where the inverse relation $\langle\xi_0\xi_0^\dagger\rangle_{J(\bar{\nu}_0)}\equiv\bar{\nu}_0$ expresses $\boldsymbol{J}$ in terms of the most convenient number expectation, and the second and third lines of Eq. (F16) have been filled in from the closed-form evaluation of Eq. (F14). Since invertible relations

$$\bar{\nu}_0 = (K_0+I)^{-1}[\bar{n}_J(0_-)+I] = [\bar{n}_J(0_+)+I](K_0+I)^{-1} \tag{F17}$$

exist, $\bar{\nu}_0$ is a convenient and acceptable independent variable to encode the initial (and final) conditions on $\bar{n}$.

On the other hand, from the original form of the generating functional (F1), and the fact that by definition of the time loop,

$$\exp\left(i\oint_0 dt\,\mathrm{Tr}[E\hat{n}]\right) = 1, \tag{F18}$$

it would seem that the proper Legendre transform should have some form like

$$\zeta_{\bar{n}} \sim \exp\left(i\oint dt\,\mathrm{Tr}[J(t)\bar{n}_J(t)]\right)\zeta_J\Big|_{J=J_{\bar{n}}}. \tag{F19}$$

In fact, the correct transform must have both real and imaginary terms, because the real term in Eq. (F16) is necessary to handle the boundary values, but not dependent at all on the history of which Eq. (F19) should be a functional.

The key to correctly separating the dependencies is counting the degrees of freedom and ambiguity in the original history n_J, and in any putative inverse function J_n. Clearly, the evolution equation (F12) cannot give dynamics to the trace of $\tilde{n}_J$, so one component of any inversion $\tilde{J}_n$ must be arbitrary except possibly at $t=0$. This freedom can be systematically parametrized by splitting

$$\tilde{J}_n(t) = \tilde{J}_\perp(t) + \lambda(t)\tilde{n}(t), \tag{F20}$$

such that

$$\mathrm{Tr}[\tilde{J}_\perp(t)\tilde{n}(t)] = 0, \tag{F21}$$

leaving $\lambda(t)$ an almost arbitrary function. There may be additional nondynamical components of $\tilde{n}_J$, such as the determinant if J is Hermitian, in which case there are additional almost arbitrary components in $\tilde{J}_\perp$. Such components could be fixed by any regular constraints (resembling gauge conditions), except that they must recover Eq. (F15) globally, to encode the finitely many degrees of freedom in $\tilde{n}(0_\pm)$.

One way to handle the global inversion is to let $\tilde{J}_n$ have a singular component,

$$\tilde{J}_n(t) = \tilde{J}_{\text{reg}}(t) + \tilde{J}_+ \delta(t - 0_+), \tag{F22}$$

imposing gauge conditions everywhere on $\tilde{J}_{\text{reg}}$ to fix $\lambda(t)$ and any other free functions, and letting the final transformation generated by $\tilde{J}_+$ produce the proper J. Using such an inversion, the correct Legendre transformation of ζ_J is

$$\zeta_{\bar{n}} \equiv \exp\left(i \oint dt\, \text{Tr}[\tilde{J}_{\text{reg}}(t)\tilde{n}(t)] \right) e^{\text{Tr}[J\bar{\nu}_0]} \zeta_J|_{J(\bar{\nu}_0)}. \tag{F23}$$

Under variations that leave initial $\bar{n}$ fixed,

$$-i\frac{\delta \ln\zeta_{\bar{n}}}{\delta\bar{n}(t)}\bigg|_{\bar{n}(0)} = J_{\text{reg}}(t) = J_{\bar{n}}(t), \tag{F24}$$

because J_{reg} differs from $J_{\bar{n}}$ only at $t=0_+$. The solutions $\bar{n}(t)$ where $J_{\text{reg}}(t)\equiv 0$ are the classical expectations $\bar{n}_{\text{cl}}(t,\bar{\nu}_0)$ under the intrinsic dynamics. Equation (F24) thus vanishes at the right solutions to define an effective action.

Meanwhile, variations of the total solution under which $\bar{n}(t)$ remains the classical stationary point with respect to its initial conditions are just those of Eq. (F16):

$$\frac{d \ln\zeta_{\bar{n}}}{d\bar{\nu}_0}\bigg|_{\bar{n}(t)=\bar{n}_{\text{cl}}(t,\bar{\nu}_0)} = J(\bar{\nu}_0). \tag{F25}$$

J vanishes on the number expectation value set by the original thermal partition function, so the effective potential has the same relation to initial conditions that the effective action has to dynamics.

The Legendre transform of a finite-temperature partition function with nontrivial dynamics thus takes the form

$$\zeta_{\bar{n}} = e^{iS_{\text{eff}}[\bar{n}]} e^{-\bar{\beta}\mathcal{F}[\bar{n}]}, \tag{F26}$$

where for the linear, Gaussian-coherent ensemble,

$$S_{\text{eff}}[\bar{n}] = \oint dt\, \text{Tr}[J_{\text{reg}}(t)\bar{n}(t)] \tag{F27}$$

and

$$\bar{\beta}\mathcal{F}[\bar{n}] = -\ln Z + \text{Tr}[K_0\bar{\nu}_0 - I - \ln(K_0\bar{\nu}_0)]. \tag{F28}$$

Though the constructions demonstrated here referred to the form of this particular ensemble, the separation between dynamics and initial values is natural, and the same form for the effective potential should arise for any theory that can be expanded in Gaussian fluctuations.

3. Relation to the Matsubara path integral

The argument of the exponential in Eq. (F26) has a natural interpretation as a single contour integral, provided by the Matsubara path-integral construction of finite-temperature
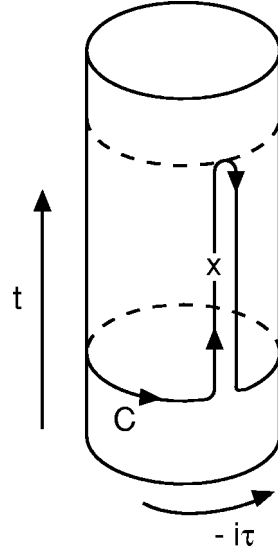

FIG. 5. Deformation of the Matsubara contour to include a segment of time-loop S matrix. Rotation $t \leftrightarrow -i\tau$ gives the signs for the action and free energy in Eq. (F23) from a single integral with contour $\mathcal{C}$, and x marks the leg on which positive-time correlations are measured.

ensembles. In the Matsubara theory, the partition function is a functional integral of fields on a periodic contour. Correlations on the contour have an analytic continuation to retarded Green's functions [39], giving the contour an interpretation as a rotation of a time axis to periodic imaginary values. If a stationary-point expansion of the functional integral were performed, the log of the partition function βF would itself be a contour integral, with contour period β, and F a density on the imaginary time axis (see Ref. [20], Sec. II A, for an example). The leading ("classical") contribution to F would be the field Hamiltonian at the stationary configuration, and the determinantal correction would, in general, be proportional to the contour length, as is generally the case in local field theories [40].

Since the effective action in Eq. (F26) is itself a time-loop integral, with starting and ending points at $t=0$, it may be regarded as an insertion into the imaginary contour integral already defining $\bar{\beta}\mathcal{F}$ (the multitemperature generalization of βF). At that point, it effectively becomes a continuation of the original Matsubara contour to the (real) time axis in a periodic, complex manifold, on which all of the Green's functions are defined, as shown in Fig. 5. The continuation of the stationary-point evaluation of the effective potential to the classical effective action is no more than the cumulation of the analytic continuations of the microscopic thermal to retarded Green's functions, so the form of the Legendre transform could have been anticipated. Further, because the contour defining a thermal trace is intrinsically closed, it must have been the time-loop S matrix, and the resulting time-loop action it generates, that embeds that trace in real time.

This derivation of classical actions from the stationary-point evaluations of Matsubara Hamiltonians was proposed in Ref. [20]. Both in that work and here, nontrivial dynamics

arose from initial conditions of heterogeneous temperature. In Ref. [20], the analytic structure of the path integral was made well defined by introducing a conformal factor on the Matsubara manifold, which was itself a dynamical field, whose fluctuations were adiabatic transformations. Here, heterogeneous temperature is built into an algebraic partition function by exploiting convenient features of Gaussian coherent ensembles. Yet adiabaticity remains central to the construction (as discussed in the text), and the form of the resulting effective potential is the same.

4. Sample forms in two dimensions

The foregoing constructions are easy to illustrate for the case K_0 2×2 and Hermitian. This case exactly describes the S and A even-state occupation numbers at $M = 1$, analyzed in the text as a refrigeration example, after the odd-state ξ components are integrated out of the Gaussian integral. Hermiticity of K_0 ensures real $\bar{n}_0$ eigenvalues, hence real observed mean number in all physical bases. Probing with Hermitian $\tilde{J}$ illustrates the gauge freedoms introduced when both the determinant and trace of $\tilde{n}_J$ are nondynamical.

The decomposition of a 2×2 Hermitian $\tilde{J}$ into its independent vector degrees of freedom is

$$\tilde{J} = \sum_{i=0}^{3} \tilde{J}_i \hat{e}_i. \tag{F29}$$

The basis matrix $\hat{e}_0 \equiv I$, while

$$\hat{e}_1 = \begin{bmatrix} & 1 \\ 1 & \end{bmatrix}, \tag{F30}$$

$$\hat{e}_2 = \begin{bmatrix} & -i \\ i & \end{bmatrix}, \tag{F31}$$

$$\hat{e}_3 = \begin{bmatrix} 1 & \\ & -1 \end{bmatrix}. \tag{F32}$$

The corresponding decomposition of the number operator is

$$\hat{\boldsymbol{n}} = \sum_{i=0}^{3} \hat{n}_i \hat{e}_i \tag{F33}$$

with the component operators having the physical identifications in the engine model

$$\hat{n}_0 = \tfrac{1}{2}(\hat{n}_x + \hat{n}_y) = \tfrac{1}{2}(\hat{n}_L + \hat{n}_R) = \tfrac{1}{2}(\hat{n}_+ + \hat{n}_-),$$

$$\hat{n}_1 = \tfrac{1}{2}(\hat{n}_L - \hat{n}_R),$$

$$\hat{n}_2 = \tfrac{1}{2}(\hat{n}_+ - \hat{n}_-),$$

$$\hat{n}_3 = \tfrac{1}{2}(\hat{n}_x - \hat{n}_y). \tag{F34}$$

The coefficients of the $\hat{e}_i$ in this representation form SO(1,3) vectors, which for the number operator may be de-

noted $\tilde{n} = (\tilde{n}_i)_{i=0,\ldots,3} \equiv (\tilde{n}_0, \vec{n})$, and for the probe $\tilde{J} = (\tilde{J}_i)_{i=0,\ldots,3} \equiv (\tilde{J}_0, \vec{J})$. The basis matrices have been chosen to give the coefficients in Eq. (F34) convenient normalizations, with the result that they take on commutation relations

$$[\hat{e}_i, \hat{e}_j] = 2\epsilon_{ijk}\hat{e}_k \tag{F35}$$

(with ϵ_{ijk} the totally antisymmetric symbol on three indices), and traces

$$\text{Tr}[\hat{e}_i \hat{e}_i] = 2, \quad \forall\, i. \tag{F36}$$

It is clear from the evolution equation (F12) that $d(\tilde{n}_0)_J/dt = 0$, and from the commutation relations (F35) that the vector $\vec{n}_J$ evolves as

$$\frac{d\vec{n}_J}{dt} = 2\vec{J} \times \vec{n}_J. \tag{F37}$$

Thus, $d(\vec{n}_J \cdot \vec{n}_J)/dt = 0$ as well.

The first gauge condition that can be put on an inversion $\tilde{J}_n$, to make up for the degree of freedom unfixed by a dynamical determinant of $\tilde{n}$, is $(\tilde{J}_0)_{\bar{n}} = 0$. The correct inversion of the remaining vector components then decomposes as

$$2\vec{J}_{\bar{n}} = \frac{1}{\vec{n} \cdot \vec{n}} \left(\vec{n} \times \frac{d\vec{n}}{dt} + \lambda \vec{n} \right), \tag{F38}$$

where the cross product is manifestly traceless with $\tilde{n}$, and the arbitrary function $\lambda(t)$ corresponds to the degree of freedom unfixed by a dynamical trace of $\tilde{n}$. λ may be gauged regularly by choosing $(\tilde{J}_3)_{\bar{n}} = 0$, for all solutions with $\tilde{n}_3 \neq 0$. (Some basis in which such a regular gauge condition applies can be chosen for any stationary solution $\tilde{n}$ on a finite t interval, since a regular solution cannot be space filling on such an interval.) The expansion in components, of the function λ selected by this gauge, is

$$\lambda = \frac{\tilde{n}_2}{\tilde{n}_3} \frac{d\tilde{n}_1}{dt} - \frac{\tilde{n}_1}{\tilde{n}_3} \frac{d\tilde{n}_2}{dt}. \tag{F39}$$

The trace which is the argument of action (F27) is very easy to evaluate in such an "axial" gauge,

$$\text{Tr}[J_{\bar{n}}(t)\bar{n}(t)] = \text{Tr}[\tilde{J}_{\bar{n}}(t)\tilde{n}(t)] = 2\vec{J}_{\bar{n}} \cdot \vec{n} = \lambda \tag{F40}$$

with the result that

$$S_{\text{eff}}[\bar{n}] = \oint dt\, \lambda(t). \tag{F41}$$

Variation of Eq. (F41) on the constraint surface $(\vec{n} \cdot \vec{n})$ constant, consistent with the gauge conditions imposed on $\tilde{J}$, may be checked to yield $d\tilde{n}/dt = 0$; $\forall\, t$. From this condition and the interaction-picture definition (F6), it follows that the classical stationary solution

$$\bar{n}_{\text{cl}}(t) = e^{iEt} \tilde{n}(0_-) e^{-iEt}. \qquad \text{(F42)}$$

Then, using the inversion of Eq. (F17),

$$\tilde{n}(0_-) = (K_0 + I - \bar{\nu}_0^{-1}) \bar{\nu}_0, \qquad \text{(F43)}$$

the variation of Eq. (F28) for the initial condition yields

$$\bar{\nu}_0 = K_0^{-1} = \tilde{n}(0_-), \qquad \text{(F44)}$$

as required.

---

[1] P. Glansdorff and I. Prigogine, *Thermodynamics of Structure, Stability, and Fluctuations* (Wiley Interscience, New York, 1971).

[2] G. Nicolis and I. Prigogine, *Self-organization in Nonequilibrium Systems: From Dissipative Structures to Order Through Fluctuations* (Wiley, New York, 1977).

[3] P. Bak, C. Tang, and K. Wiesenfeld, Phys. Rev. A **38**, 364 (1988).

[4] T. Hwa, and M. Kardar, Phys. Rev. **45**, 7002 (1992).

[5] J.M. Carlson and J.S. Langer, Phys. Rev. Lett. **62**, 2632 (1989).

[6] D.S. Fisher, in *Nonlinearity in Condensed Matter*, Proceedings of the Sixth Annual Conference, Center for Nonlinear Studies, edited by A.R. Bishop, D.K. Campbell, P. Kumar, and S.E. Trullinger (Springer-Verlag, Berlin, 1986); Phys. Rev. Lett. **50**, 1486 (1983).

[7] B. Bhushan, J.N. Israelachvili, and U. Landman, Nature (London) **374**, 607 (1995).

[8] G. Nicolis and I. Prigogine, *Exploring Complexity* (Freeman, New York, 1989).

[9] P. Bak, *How Nature Works* (Copernicus, New York, 1996).

[10] J.M. Carlson and J. Doyle, Phys. Rev. Lett. **84**, 2529 (2000).

[11] R.M. D'Sousa and N.H. Margolus, Phys. Rev. E **60**, 264 (1999).

[12] E. Fermi, *Thermodynamics* (Dover, New York, 1956).

[13] K. Huang, *Statistical Mechanics* (Wiley, New York, 1987).

[14] M. Gell-Mann and S. Lloyd, Complexity **2**, 44 (1996).

[15] W.H. Zurek, Nature (London) **341**, 119 (1989); Phys. Rev. A **40**, 4731 (1989).

[16] S. Lloyd, in *Elementary Particles and the Universe: Essays in Honor of Murray Gell-Mann*, edited by J.H. Schwarz (Cambridge University Press, Cambridge, 1991).

[17] J.P. Crutchfield and K. Young, Phys. Rev. Lett. **63**, 105 (1989).

[18] C.R. Shalizi and C. Moore, e-print cond-mat/0303625.

[19] E.T. Jaynes, *Papers on Probability, Statistics and Statistical Physics*, edited by R.D. Rosenkrantz (Reidel, Dordrecht, Holland, 1983).

[20] E. Smith, Phys. Rev. E **58**, 2818 (1998).

[21] E. Smith, Phys. Rev. E **60**, 3633 (1999).

[22] P. Ceperley, J. Acoust. Soc. Am. **66**, 1508 (1979); **72**, 1688 (1982); **77**, 1239 (1985). For a tutorial overview of the closely related irreversible thermoacoustic engines, see Ref. [33].

[23] S. Coleman, *Aspects of Symmetry* (Cambridge University Press, New York, 1985), Chap. 5.

[24] S. Weinberg, *The Quantum Theory of Fields* (Cambridge University Press, New York, 1996), Vol. I. (Most general theory consistent with the symmetries.)

[25] D.C. Mattis and M.L. Glasser, Rev. Mod. Phys. **70**, 979 (1998).

[26] G.L. Eyink, Phys. Rev. E **54**, 3419 (1996).

[27] A.J. Organ, *Thermodynamics and Gas Dynamics of the Stirling Cycle Machine* (Cambridge University Press, New York, 1992).

[28] A.V. Andreev, V.I. Emel'ianov, and IU.A. Il'inskii, *Cooperative Effects in Optics: Superradiance and Phase Transitions* (Institute of Physics Pub., Philadelphia, 1993).

[29] G. Backhaus and G.W. Swift, Nature (London) **399**, 335-8 (1999); J. Acoust. Soc. Am. **107**, 3148 (2000).

[30] L. Onsager and S. Machlup, Phys. Rev. **91**, 1505 (1953).

[31] M. Mareschal, M. Malek Mansour, A. Puhl, and E. Kestemont, Phys. Rev. Lett. **61**, 2550 (1988).

[32] A. Zhabotinsky, Biofizika **9**, 306 (1964).

[33] G.W. Swift, J. Acoust. Soc. Am. **84**, 1145 (1988).

[34] L. Stryer, *Biochemistry* (Freeman, San Francisco, 1981).

[35] P. Jizba and T. Arimitsu, in *Nonequilibrium and Nonlinear Dynamics in Nuclear and Other Finite Systems*, edited by Zhuxia Li, Ke Wu, Xizhen Wu, Enguang Zhao, and Fumihiko Sakata, AIP Conf. Proc. No. 597 (AIP, Melville, NY, 2001), pp. 341–348.

[36] E.F. Schrödinger, *What is Life?: The Physical Aspect of the Living Cell* (Cambridge University Press, New York, 1992).

[37] H.J. Morowitz, *Beginnings of Cellular Life; Metabolism Recapitulates Biogenesis* (Yale University Press, New Haven, 1992).

[38] D. Park, *Introduction to the Quantum Theory*, 2nd ed. (McGraw-Hill, New York, 1974).

[39] G.D. Mahan, *Many-Particle Physics*, 2nd ed. (Plenum, New York, 1990), Sec. 2.9.

[40] S. Coleman, *Aspects of Symmetry* (Ref. [23]), Chap. 7, and its App. 1.

[41] This idealization glosses over many real differences in polymerization processes, such as the frequent use of prephosphorylation of monomers as the energy-injecting step. The current idealization is chosen to make the correspondence visible with the fewest qualifications and contortions.

[42] Of which animal ears are good examples. The evolutionary robustness of the Fourier-transform capabilities of ears may be evidence for the ubiquity of frequency-diagonal state variables in the natural world.